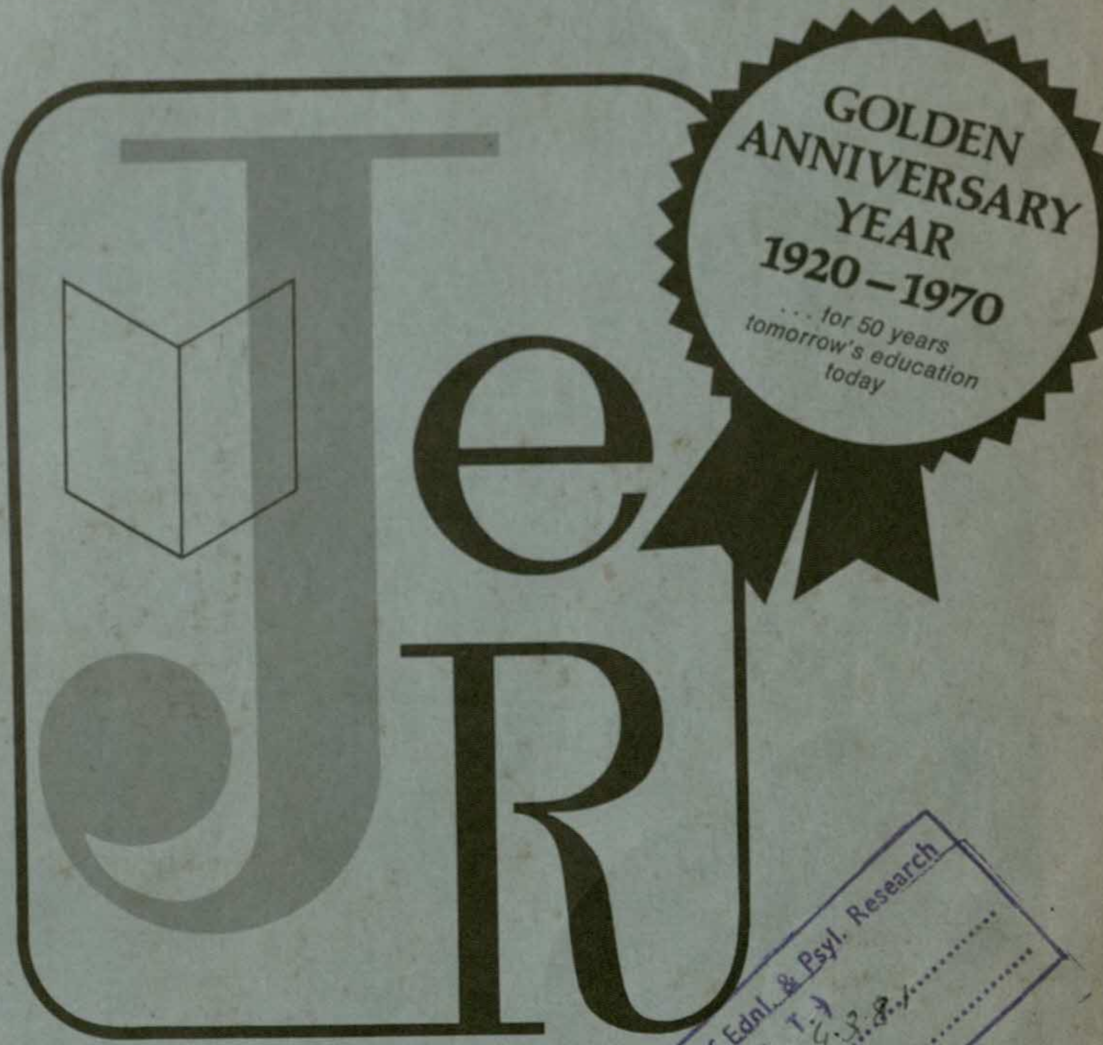
K.S.
Docket
issue
the library
9/12/7

The Journal of Educational Research

Dedicated to the Study of Education



GOLDEN ANNIVERSARY YEAR
1920 – 1970
. . . for 50 years tomorrow's education today

Volume 64 ~~SEPTEMBER~~ 1970 -71 ~~Number 1~~

Bureau of Ednl. & Psyl. Research
(S. C. E. R. T.)
Date 4.3.8.1
Acc. 5664

IN THIS ISSUE: Innovation in Education

Policy and Perspective

by **LINDLEY J. STILES**

Professor of Education for Interdisciplinary Studies, Sociology, and Political Science, Northwestern University

Nutrition and Early Mental Development

GUEST CONTRIBUTOR
Leonardo Sinisterra, M. D.*
Professor of Nutrition, School of Medicine
University of Cali
Colombia, South America

DURING THE past decade research in different disciplines and in different parts of the world has suggested a relationship between the way children are fed and their mental and behavioral development. Undernutrition produces physical underdevelopment (small size, low body weight), and now it is known to be responsible also for mental underdevelopment. Implications of these findings are of significance to educators as well as to medical doctors. Approximately two-thirds of the inhabitants of the world suffer from chronic malnutrition (undernutrition): these are the people living in the so-called underdeveloped countries. In the affluent, developed nations of the northern hemisphere, migration of the rural populations to the urban areas has produced the slums where undernutrition is also highly prevalent. As a consequence, undernutrition, especially harmful in the early infant years, is a positive threat to close to four-fifths of the inhabitants of the world—and this is happening in the last third of the twentieth century.

The biological impact of undernutrition is well known. Growth of human tissues (as well as any animal tissue), is the final result of the synthesis of protein by the subject, starting from the protein he or she has received in the diet. Protein of the diet has to have certain characteristics of biological quality if the subject is to grow adequately as the result of its use. Biological quality of proteins derives from the relative content of each of the chemical units called amino acids. Of the twenty-two amino acids, eight are called indispensables because they cannot be synthesized by the human organism and as such they have to be provided by the protein of the diet. The rest are dispensable in that they can usually be produced within the human organism.

Of significance is the fact that only animal proteins have in their structure all the needed amino acids: both dispensable and indispensable ones. This is why animal proteins are said to have a higher biological quality than vegetable proteins. Unfortunately, the adequate use of animal proteins (the protein of milk, meat, eggs and other animal products) in the diet is the practice in only a few countries of the world. In most of the countries, prices are too high which produces a limited market, despite the need. The vicious circle of low production and low use that results is the basis for the "human underdevelopment" that maintains.

Most of the world's population survives on vegetable, starchy foods which provide very little protein and that available is of poor biological quality. Southeast Asia's rice and Latin America's corn are low in Lysine and Tryptophane—two indispensable amino acids. Africa's yams, sweet potatoes, and cassava are almost devoid of protein; they provide only sugar. The diets of the slum dwellers in more affluent nations consists, primarily, of soft drinks, fried potatoes, bread and other foods of low price and low biological quality.

The impact of poor nutrition on mental development is most critical during the early years of life: from conception to 6 years of age. It is during this period that the brain achieves 90 percent of its growth. Undernourishment during these early years cannot be corrected after the child reaches school age. Hence, efforts by schools (in hot lunch programs, for example) come too late to prevent permanent mental damage. A concomitant with the poor diet that children often receive during their critical preschool years is a dull psychological and social environment. They have nobody to play with them or to give them in other ways the attention that stimulates mental responses and development. One observes that they stand around most of the time by themselves—undernourished, understimulated, and underdeveloped.

With the deficiencies in biological and emotional environment that they confront, it should be no surprise that these children will have difficulty in school. Because their slow pace of learning is due to underdeveloped mental capacities, educational programs—however well designed they may be—cannot make up for the deficiencies. The children will fail to learn in school and fail to contribute to society.

The underdeveloped child is a victim of environmental circumstances. Educators, along with health officials and medical doctors, confront the problem of changing these circumstances to give all children a chance for normal development in the preschool years. To change the nutritional, intellectual, and emotional forces that children confront, help must be provided for the family, and particularly for the mother who carries the major responsibility for rearing children.

Mothers, as well as the total society, unfortunately, are almost completely ignorant of sound principles of adequate child nutrition. Nor do they have an understanding of modern educational and psychological child-rearing practices. It is at this point that the school has a key role to play. It can provide, through theoretical teaching and practical demonstrations to mothers and the community in general, the knowledge needed to provide proper nutrition as well as emotional and psychological stimulation to preschool children. Such educational help to mothers by schools is easily justified since mothers have under their care and responsibility the most precious asset of the country and the world—the future citizens.

**Dr. Sinisterra was a visiting scholar at Northwestern University during the 1969-70 academic year. His research on nutrition for preschool children is well known throughout the world.*

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor Curriculum and Instruction, The University of Wisconsin, Madison.

PHILIP LAMBERT, Professor of Educational Psychology and Administration, The School of Education, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1970

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Assistant Professor of Special Education, Yeshiva University, New York.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

NORMAN R. STEWART, Associate Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English and Education, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLASS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

CARLETON W. WASHBURNE, Director Emeritus of Teacher Education and the Division of Graduate Studies, Brooklyn College, 2248 Kent Street, Okemos, Michigan 48864.

Bureau Econl. Psy. Research
File N L.L.
Date 9.12.70
Diary No. 179

50th Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **SEPTEMBER 1970** **NUMBER 1**

ARTICLES

Toward a National Strategy for Environmental Education—*Clay Schoenfeld* 3

Levels of Cheating Under Conditions of Informative Appeal to Honesty, Public Affirmation of Value, and Threats of Punishment—*Constance T. Fischer* 12

Credibility of Information for Educational Innovation—*Frank B. Murray* 17

Social Maturity in the Prediction of Academic Achievement—*Thomas Arena* 21

Factors Relating to the Innovativeness of School Systems— *Leo R. Hilfiker* 23

Conceptual Pre-Structuring for Detailed Verbal Passages—*Barton B. Proger, Raymond G. Taylor, Jr., Lester Mann, John M. Coulson, Robert J. Bayuk, Jr.* 28

Use of the Canonical Correlation Analysis in an Investigation of Pupil's Rate of Learning in School—*Margaret C. Wang* 35

DEPARTMENTS

Among Those Present 48

Book Reviews—*Robert E. Clasen* 20

Field News—*Willard E. North* 46

Forthcoming Articles 16

It Says in The JER 2

Research Brief—*Edward P. Merryman* 34

(The Journal of Educational Research is indexed in Abstr.S.W., CSPA, Current Contents, Ed. Adm. Abst., Educ. Ind., M.H.R.Ind., PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1970 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by **DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI**, President. Prof. **WILSON B. THIEDE**, Vice President and Publisher. Prof. **CLARENCE A. SCHOENFELD**, Assistant to the Publisher. **ARNOLD CAUCUTT**, Treasurer and Business Manager. **NORMA COURNOW CAMP**, Supervisor of Editorial Services.

It Says in The JER...

A NATIONAL ENVIRONMENTAL EDUCATION PROGRAM is envisaged by Schoenfeld. He sees environmental organization beginning at the Presidential level, going to the Cabinet, Regional, State, and local levels. The program would be multidisciplinary in structure and multi-process in programs (p. 3).

ELEMENTARY SCHOOL STUDENTS were given an opportunity to cheat in Fischer's study. She found that public affirmation of a belief in not cheating is the preferable classroom method of minimizing cheating (p. 12).

EDUCATORS AND RESEARCHERS are in disagreement about the importance of standardized test results but agree that the educational research journal is the most credible source of educational information. Murray compared the rankings of educational psychology professors and junior high teachers on eight kinds of information in terms of how important each would be in their decision to adopt a particular educational innovation and found that they failed to correlate significantly (p. 17).

THE SOCIAL MATURITY OF NORMAL CHILDREN, as measured by the Vineland Social Maturity Scale, was of no significance in predicting academic success, concludes Arena after studying fifty-two randomly selected fourth-graders. Likewise, the addition of social maturity scores did not significantly increase the predictive value of intelligence test scores (p. 21).

SCHOOL SYSTEM SELF-STUDY and awareness of the degree of openness, trust, etc., perceived to exist throughout the system could be of assistance to the personnel of the school system in knowing more about themselves and about the potentialities of the organization, says Hilfiker. If education is to become a part of the social change taking place, more knowledge is needed regarding effective ways and means of instituting changes and innovations in the school system (p. 23).

SECONDARY STUDENTS' understanding of key points of a printed passage can be improved through methods of conceptual organization. Proger and others found that the paragraph abstract and sentence outline methods gave significant facilitation to the groups using them (p. 28).

RATE OF LEARNING, Wang concluded, is specific to a given task and is not a general factor characterizing elementary students' performance in all learning situations (p. 35).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Editorial/Abstract

Toward a National Strategy for Environmental Education

◆ This paper looks toward Federal legislation providing concepts, structure, and funding for a major coordinated educational effort at the elementary–secondary, college–university, and adult-community levels to inculcate the mass conservation literacy necessary to support the public and private quest for environmental quality.

We are talking, then, about environmental education.

The term "environmental education" is used here in the broadest sense of its meaning—development of an awareness of the natural physical environment, a recognition of the interdependence among natural resources and of man's complete dependence upon them, an understanding of intelligent resource conservation and improvement of the physical environment. In essence, environmental education views resources as a community of which man is a part, not as a commodity which man is to exploit. It seeks to lay a basis for action in the public interest, to elucidate the choices in land and water use and relate them to general values and social objectives, to provide integrated approaches to environmental management consistent with ecological principles, economic facts, esthetic insights, and ethical dimensions.

Some might say the term has sprung into being merely at the whim of phrasemakers to lend a charismatic quality to the matter with which it is associated. On the contrary, the term is coming into use to satisfy the very real need of people to describe, if not a new program or set of programs, at least a new way of looking at a variety of old programs, their relationships, and their potential contributions. Too, by becoming associated with particular pressure groups and specific methods, the term conservation education has tended to lose some of its viability as an approach to an integrated man–land ethic. Conservation education lacks compelling theories of content and methodology; its concepts and materials are rooted in an agrarian era; it lacks practitioners of adequate background and scope; it is oriented to special "users" rather than to the environment itself. It has no clear-cut coordination, impetus, or funding.

Hence the search for a planning strategy for the new environmental education in America.

While this article does not look at the field of environmental education strictly from a research base, it deals with a problem so important to the United States that it was felt to be a significant contribution to the *Journal's* Golden Anniversary Series. It emphasizes particularly the need for regional Environmental Education Centers at selected universities: The proposed Environmental Education Centers would put together interdisciplinary teams of experts in environmental management and charge them with collecting, collating, and disseminating practical guidelines to community organization and action in resource conservation. Each Center would also bring to the campus for a year of advanced study and practical work promising "change agents" from community and regional levels in such specialties as law, planning, public administration, resource management, education, and communications. Further, each Center would engage in research in adult education and communications as they relate to the quest for environmental quality.

The author is Joint Professor of Journalism and Wildlife Ecology and Chairman of the Center for Environmental Communications and Education Studies at The University of Wisconsin—Madison. His main philosophy:

"Strategically, we must put our trust in an educational system which protects diversity—diverse species, diverse attitudes, diverse adaptations, diverse experiments, diverse thrusts. We must develop new structures, new strategies, new techniques. We must test and revise until we have developed a culture which recognizes man's interdependence with his environment and with all of life, and his responsibility for maintaining that environment in a condition fit for life and fit for living."

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Toward a National Strategy for Environmental Education

CLAY SCHOENFELD
Editor, The Journal of **Environmental Education**

IN the past 50 years in which *The Journal of Educational Research* has been recording the growing edge of American education, our schools and colleges have been asked to support a series of national goals—the melting pot, prosperity, depression pump-priming, national defense, the conquest of space, and so on. Today's mission is nothing less than human survival.

We are figuratively and literally sick and tired of "the mis-development of America," a degradation of our environment that diminishes daily the quality of the human experience: water pollution; air pollution; soil erosion; forest, range, and wetland deterioration; waning wildlife; urban sprawl; preempted open spaces; vanishing wilderness; landscapes scarred by highways, litter, noise, and blight—a not-so-quiet crisis of decreasing beauty and increasing contamination that threatens not only the pursuit of happiness but life itself (16).

To do something about environmental conservation, redevelopment, and maintenance requires a sense of husbandry, a sense of responsibility on the part of every American citizen—that man-land ethic or "ecological conscience" which Aldo Leopold bespoke. It is unthinking people who pollute the environment; it is thinking people who can effect a "new conservation": to restore as well as to protect, to bring beauty to the cities as well as to keep it in the countryside, to handle the waste products of technology as well as the waste of natural resources, to halt the massive deterioration in the American environment, to husband those resources and amenities which are inextricably linked not only to economic prosperity but to the inner prosperity of the human spirit (8).

We are talking, then, about environmental education: a recognition by man of his interdependence with his environment and all of life, and his responsibility for developing a culture which maintains that relationship through policies and practices necessary to secure the future of an environment fit for life and fit for living (2).

Many ask, what is the difference, if any, between the new "environmental education" and the older "conservation education?" The newer term attempts to do a more precise and at the same time a more comprehensive job of describing our ecological efforts to come to grips with the degradation of man's interlaced surroundings.

In another paper (17) I have gone to some lengths to delineate the critical differences between the old "conservation" and the new "environmentalism." While some differences may be more apparent than real, others are quite distinctive. They may be summarized as follows (11):

In terms of its scope, the new environmentalism attempts to be all-encompassing. Whereas yesterday we tended to treat soil conservation, water conservation, forest conservation, wildlife conservation, and so on as separate units, today we try to understand and explain the ecological unity of all man–land relationships. In terms of its focus, then, the new environmentalism is man-centered. That is, our primary concern has shifted from the survival of remnant redwoods and raptors to the survival of nothing less than the human species itself. At the same time we are not so much concerned about quantities of natural resources as we are about the quality of the human experience.

In terms of its locus, while the old conservation conjured up images of open country, the new environmentalism incorporates the pressing problems of the city. In terms of its emotional underpinnings, the new environmentalism is based more on fear for man's tomorrow than on a love for nature's yesterday. Thus today's "preservationist" is not a lover of wilderness; he is one who fears the four horsemen of "conquest, slaughter, famine, and death." In terms of its political alliances, the old conservation was linked to such orthodox causes as depression pump-priming, national defense, and outdoor recreation; the new environmentalism, on the other hand, encompasses the hitherto unmentionable demands of the neo-Malthusians for population control.

It is in its basic cultural orientation, however, that the new environmentalism differs most strikingly from its antecedent—conservation. The latter, in the words of one patron saint, stood clearly for economic development, for the infinite goodness of American "progress" (7). But environmentalism reflects a growing suspicion that bigger is not necessarily better, slower can be faster, and less can be more.

If anything surely marks this revolutionary nature of both the rise and rationale of the new environmentalism, it would be the recent words of a Republican President of the United States, telling us that "wealth and happiness are not the same thing," that now is the time to "make our peace with nature," and that we must "measure success or failure by new criteria" (6). Not even FDR would have said that!

Hence the search for a planning strategy for the new environmental education in America. If a national effort in environmental education is to be strengthened, we must (1) identify the audiences, (2) suggest methods, (3) outline an organization, and (4) propose a timetable. In terms of a timetable, this paper proposes that we proceed simultaneously on a variety of fronts. In terms of an organization, this paper proposes that we build "critical masses" of environmental education at the federal, state, and local levels. In terms of methods, this paper proposes that we start with what we have and invest heavily in research that will reveal optimum tactics. In terms of audiences, this paper suggests three: elementary and secondary school pupils and their teachers, college and university students and their professors, and adults in their roles as leaders of community action and as citizens in general.

Adult Education and Community Service

The older Smith–Lever Act and the newer Title I of the Higher Education Act provide entres to adult education in environmental conservation through university extension. But they are not enough in scope and intensity. We need to establish on appropriate university campuses Environmental Education Centers that will focus on the problem of translating resource management policies and plans into action on the landscape through (a) the collation of information about, and the preparation of practical materials on, restoring the quality of the environment, (b) informal instruction and technical assistance carried out directly to local governments, regional instrumentalities, and citizen groups, (c) the refresher education of key practitioners brought back to the campus for work in natural resource policy implementation, and (d) research in adult education theory and practice. Growing national programs of environmental management depend increasingly on state and local initiative and responsibility. "The problem of the states" as *Life* magazine said in its editorial of Nov. 4, 1966, "is not lack of power or opportunity or even solely of money; it is a shortage of competent public servants." The proposed Environmental Education Centers would tackle the problem directly by upgrading such key community leaders as planners, lawyers, resource specialists, adult educators, communicators, teachers, and public administrators, and by reinforcing them with improved educational materials and professional consultation. The Centers would also be concerned with adult education and communications research.

The Problem

Why do we continue to have serious trouble translating federal intentions and state plans into timely, sound action on the land? One answer would certainly seem to be it is only at the local and regional level that public policies can be translated into public and private practices, and it is here at the intrastate level where the forces of wise resource management continue to be ill-equipped to deal with the forces of exploitation. The inefficiency with which public agencies and private citizens go about performing the socially essential tasks of environmental housekeeping stem largely from the fact that the technical and organizational skills available to the land conserver or rational planner are normally inferior to those available to the land exploiter.

While we do not of course yet understand all the scientific facts and societal values attendant to environmental quality control, what we *do* know is not being applied on a scale commensurate with the present pace of environmental pollution. At the local and regional level, where most of the decisions affecting the quality of the environment are made, we must address ourselves firmly to laying a basis for action by elucidating the choices in land and water use, relating them to general values and social objectives, instilling in people a desire for constructive change, and providing practical guidelines that encompass integrated rather than unilateral approaches.

The Approach

The urgency of need will not be met through educational processes of normal pace and dispersion. What is needed now, and for some time to come, is a steady stream of specific skills and resources rifled to the local and regional firing line. We need local leaders equipped with an understanding of the interrelationships involved between their callings and total environmental management, and with a knowledge of the "tools of the trade" in energizing land and water use controls.

To begin to develop and equip a cadre of local resource management leaders, it will be helpful to concentrate on those individuals at the local and regional level who typically do or can play the role of "change agents" in conservation. Four such categories of key personnel can be identified: (a) lawyers, (b) planners and administrators, (c) field resource technicians, and (d) communicators and educators.

To retrain all such change agents and to equip them and their clientele with effective education-for-action materials and services will be a role of the Environmental Education Centers. The Centers will also perform related research focused on improving their concepts and techniques. Initially an Environmental Education Center would assemble from the university and elsewhere as necessary a staff representing overall competence in environmental problems, to include people trained in biology, design, soils, economics, geology, engineering, ecology, sociology, public administration, journalism, education, law, and other environmental disciplines applicable to land-use issues. Their extension function would be the collection and collation of what is known, what has been done, what is being done, and what might be done; the preparation of educational materials, utilizing a variety of media; and the dissemination of information and technical counsel to local governments and citizen groups through institutes, audio and visual media, and consultation. The team would be encouraged to work with maximum speed and practicality, yet with intellectual breadth so that the environmental caretakers on the receiving end develop a sound view of the interrelatedness of decisions affecting the environment and of the immense complexity of human needs that must be reconciled and met.

At the same time, the Center would provide fellowships for practicing change agents to come to the campus for a year of study focused on the practical problems faced by those trying to plan, acquire, maintain, and manage lands and waters for public purpose. These professionals would take regular courses and special interdisciplinary seminars, and will also perform research duties in the Center. Appropriate degrees could be awarded to qualified personnel (14).

College and University Education

At the college and university level we are at the point where education in the physical sciences was 125 years ago, when scientists were debating whether a course in physics would be useful and whether it would not be just as feasible to continue work in natural philosophy. We are at the point where a course in environmental management needs organization at the university and college level. Such a program should deal broadly with the development of a conservation literacy, aimed at contributing to a general knowledge of resource management science, technology, history, organization, and philosophy, and to a basic understanding of current issues and problems calling for intelligent citizen decisions. We are turning out too many engineers that are ecologically ignorant and esthetically insensitive, too many biologists that are sociologically and economically naive. Our major goal should be to help students develop broad ideas and concepts useful in approaching resource problems. If anything, the program should emphasize that resource management issues are not simple and the answers not easy; yet that in the absence of a concensus on definitions and viewpoints, the educated man has a particular obligation to confront the problems involved with all the geobiotic facts and ethical considerations he can bring to bear (10).

The Problem

The university community today is not indifferent to the agonies of the larger society. Almost all academic people are aware and concerned observers of the nation. Some are active as technical advisors and as sources of ideas for the larger community.

All are willing critics but, for the most part, critics in the nonparticipant tradition of the husband visiting a millinery shop with his wife: "I like it"; "I don't like it." In adopting such a role, the university is limited to reacting to the alternatives that time and circumstances place before it.

That is not enough. We need in the university community a focused, systematic, responsible, even aggressive concern for the manner in which society is evolving—a concern for its values and problems, and the provision of strategies appropriate to clarify those values and to solve those problems. We need men who are seeking new solutions and actively helping us toward those solutions. We need designers of the future. We need to be told how to build a better society, and how to get from here to there. Most of all, we need help in the difficult business of changing institutions.

Consider our most grievous domestic problems —the cluster of interlocking problems of environmental degradation. One would like to think that the universities have been the primary source of intellectual stimulation and enlightenment on these issues. One would like to think that university research on these matters had laid the basis for significant action. One would like to think that University people had played a key role in formulating the public policy alternatives, and in suggesting the factual or value considerations involved in each alternative.

Unfortunately, this is far from the truth. There are brilliant and effective members of the academic world who have contributed to our approach to these problems. But generally speaking, when one moves from the arena of scientific and technical problems to those problems involving change in human institutions, one cannot say that the universities are a significant intellectual base for the main attack. In fact, a good many university people whose fields should give them a legitimate interest in these matters barely understand what the relevant problems are. Many are debating policy alternatives left behind 5 years ago. Few are planning the kind of research that would sharpen policy alternatives (5).

The conservation, redevelopment, and maintenance of environmental quality requires more managers or operators better equipped to apply present and emerging knowledge. At least three broad types of professionals can be identified. First, there are the resource technicians like park managers, geologists, ecologists, meteorologists, foresters, geographers, economists, sociologists, architects, hydrologists, game managers, engineers, and so on. Not only must they be well grounded in their specialty if they are to be immediately employable, but if they are to be promotable, they must be able to relate their discipline or function to that of other professionals and to the larger questions of human values and environmental policy. Second, there is the need for the broad resource generalist, perhaps the rarest of all species, who can deal effectively and creatively with whole policy issues in all of their complexity in a staff or executive role. Third, there must be "change agents" equipped with an understanding of the interrelationships involved between their callings and total environmental management, and with a knowledge of the tools of the trade in energizing land and water use controls—lawyers with pertinent legal skills, designers and planners with a grasp of resource policy implementation, communicators and educators equipped to interpret resource problems in such a way as to achieve concensus rather than conflict.

In the production of new professionals, are we talking about 4-year, 5-year, or 6-year programs? Do we start with generalists and make specialists out of some, or do we start with specialists and make generalists out of some? Do we modify, expand, broaden, intensify the curricula of existing departments and schools, or do we invent custom configurations? These are some of the questions presently agitating the university as it faces the training of environmental technicians, public policy formulators, and change agents. Whatever the type or level of professional environmental management education, the product should be specialized enough to be viable in field operations and broad enough to appreciate the complicated biotic and human phenomena over which he presides (13).

The Approach

About how best to involve the university in a broad strategic campaign against the pervasive degradation of the American environment, there is considerable discussion and experimentation. Because goals vary, institutions vary, traditions vary, perceptions of problems vary, and terms are imprecise, there is confusion of counsel. The trouble is that the problems of environmental quality defy the traditional pattern of the division of labor within the university which arranges knowledge into a series of subject matter areas concerned with segments of the physical, biological, and social sciences and the humanities. Conflicts among competing resource demands are frequent and intense, and it is becoming increasingly difficult to draw any clear quantitative or qualitative lines among the problems. Each specific resource use raises a congerie of questions about man's relationship to his total environment, and the choice of alternatives among competing resource uses involves legal, economic, social, administrative, technological, esthetic, ecological, and ethical considerations which in turn span manifold governmental agency and university department lines (4).

What is needed are incentive grants that will (a) encourage universities to develop integrated courses and programs which introduce the student to appropriate economic and ecological facts and esthetic dimensions, which lay a basis for action by elucidating the choices in land and water use, relating them to general values and social objectives, instilling a desire for constructive change, and suggesting some practical guidelines characterized by multidisciplinary approaches, and (b) support undergraduate and graduate students willing to break out of relatively narrow confines of particular departments and disciplines and prepare themselves for management careers in resource policy implementation broadly conceived and oriented.

The goal here is not to produce generalists with no specialty but specialists with breadth. Because universities and colleges differ markedly with respect to background, posture, and goals, we can expect to see developed a wide range of approaches to environmentally educated managers. In due time a national commission should examine the emerging programs and attempt to identify those with the most promising characteristics.

Elementary and Secondary Education

It is clear that the content of environmental education is qualitatively different than the content of the older conservation education. The confirming research was conducted by Dr. Robert Roth in his PhD thesis under the supervision of Prof. Milton Pella of The University of Wisconsin School of Education (9). It is hard to overemphasize the contribution of this research to meaningful environmental education. Greatly simplified, Roth's methodology was as follows: He reviewed the literature for environmental concepts. Then he interviewed eighty University of Wisconsin scholars interested or actively engaged in conservation and/or environmental education. They represented forty disciplines, including the sciences, humanities, and social studies. The concepts finally identified then were rated for relevancy to environmental education by a national panel of 350 scholars corresponding to the same forty disciplines as the University scholars and representing twenty-four universities across the country. The results were computer-analyzed, yielding a ranked order of 112 concepts from the most to the least important.

Upon further analysis of the 112 concepts, several things were especially revealing. First, a count of the words most frequently used clearly showed the breadth, or interdisciplinary nature, demanded in environmental management. Consider how broad a spectrum is covered by some of the most frequently used words in these key concepts: environment, man, populations, resources, economic, social, culture, individual, life, needs, values, long-range, political, public, quality, and society. The point is clear. Environmental management is NOT simply contour plowing, white-tail deer management, and life cycles of plants. That is, it is not the narrow focus traditionally labeled conservation. Instead, environmental management is of the broadest scope in that it requires an understanding of man and his total relationship to his environment.

The second startling revelation from analysis of these key environmental concepts is that forty-four of the fifty most important concepts, as ranked by the computer, can and should be introduced in the kindergarten through sixth-grade curriculum. The grade level determination was based upon the evaluation of 120 kindergarten through twelfth-grade teachers. Obviously, however, these concepts once introduced must extend throughout the student's education. It is manifest that educational curricula have not discussed man's relationship to his total environment in terms of energy flow, values, cultural, social, political, legal, and long-range quality implications. Nor have we provided any significant environmental instruction at the kindergarten through sixth-grade level. These educational gaps have resulted in a nation of socio-ecologic illiterates committing an unending series of ecological atrocities with little thought of long-term effect. Furthermore, these gaps have contributed to an almost total lack of communication among society's present decision-makers, and insufficient pressure from the public for broad environmental action programs (1).

The Problem

Some of the more pertinent problems that lie in the path of a successful national effort directed toward environmental education for youth through instructional programs (K-12) are a general lack of: (1) a coherent philosophy of environmental education among leaders in the field of conservation education, (2) teacher interest and background in environmental education, (3) school administrators dedicated to environmental educational programs for school systems, (4) existing programs that focus on environmental education, (5) well-conceived instructional material directed toward environmental education, (6) textbook orientation to environmental education, (7) individuals trained to serve as environmental educational consultants for school systems, (8) collegiate programs that provide adequate training in environmental education, (9) citizen concern in environmental education, (10) national, state, and local leadership in environmental education, (11) coordination among private and public conservation organizations, (12) a continuing information system directed toward environmental education (18).

Principal Approaches

A National Commission. Stapp recommends that a National Environmental Eductional Task Force, aided by a competent staff, be established to create a "National Environmental Educational Instructional Program" that could be integrated into a school's existing curriculum (K-12).

In developing a national program for school systems, early attention should be directed to the development of the program's philosophy, objectives, and structure. A first step in obtaining this information would be to identify and examine the philosophy, objectives, and structure of existing environmental educational programs. This information combined with ideas contributed by the task force should provide the national program with a valuable and meaningful philosophy and set of objectives. The second major step in the development of a national program would be to develop curricular material. The third major step would be to provide guidelines for the kinds of services that an environmental educational consultant could provide to a school system. A fourth major step in the development of a national program would be to develop guidelines for regional workshops for teachers and administrators. The workshops would be directed toward assisting participants to obtain a fuller understanding of environmental resource problems and the role of schools in providing leadership and instructional program in the area of environmental education.

U. S. Office of Education Involvement. Believing that it is the responsibility of the Federal Government to provide effective leadership in environmental education, a powerful pressure group has suggested the following actions be taken by the Office of Education, U. S. Department of Health, Education, and Welfare:

1. Assign specific responsibility and authority for recommending policy and directing the work of the Office of Education in the field of conservation education to one unit in the office of the Commissioner of Education.
2. Establish and maintain a highly qualified staff of specialists (consultants) in conservation education that is representative of the social, biological, and physical sciences.
3. Establish and support a continuing National Advisory Committee on Environmental Education as an advisory body to the U. S. Office of Education.
4. Provide support for in-depth research in conservation education, at the regional educational research laboratories operated by the Office of Education and elsewhere and through grants.
5. Give emphasis, through programs now in operation and in other ways, to helping state departments of education improve their capabilities to provide schools with expert counsel and guidance on integration of natural resource conservation concepts into all areas of the curriculum.
6. Provide specifically for support at colleges and universities of organized institutes and other forms of education and training in the field of conservation education for teachers (3).

Recapitulation

This paper envisages a National Environmental Education Program which will make organizational and fiscal provisions for the following main thrusts:

1. At the Presidential level, a National Commission on Environmental Education, charged principally with developing integrated concepts, methodologies, and materials.
2. At the Cabinet level, an integrated National Environmental Education Center, lodged in the U. S. Office of Education but drawing support from all appropriate agencies.
3. At the Regional level, integrated Environmental Education Centers charged with instruction, research, and technical services.
4. At the State level, Environmental Education Coordinators in each Chief State School Officer headquarters.
5. At the Local level, coordinators, curricula, committees, and sites to translate the national effort into action in the school and on the landscape.

It is essential that the Program be multidisciplinary in its posture and multi-process in its programs; unilateral approaches to environmental management have caused many of our present problems.

This paper has already suggested what might be meant by multi-process; namely, a program concerned with the production of new knowledge and new knowledge-seekers, of more and better resource managers, of citizenship education, and of technical counseling and services. What do we mean by multidisciplinary? We mean we are concerned with the total environment of man: its social, cultural, economic, and esthetic, as well as its physical and biological, aspects. To seek environmental quality requires both an understanding of human needs and the needs of a healthy natural environment. The development and management of environmental quality requires contributions by all the arts, sciences, and professions. The end is to bring conflicting forces into functional relationships in an order in which human impact does not needlessly destroy environmental quality and where environmental quality contributes to more fruitful human life, liberty, and the pursuit of happiness. While we recognize the essential importance of strengthening existing disciplines, the essential nature of environmental education looks toward research, teaching, and extension configurations that transcend traditional lines of endeavor and are concerned with the wholeness of the relationship between man and his surroundings.

The rationale for a National Environmental Education Program is simple: it is unthinking people who pollute the environment, and it is thinking people who can bring about environmental conservation, redevelopment, and maintenance. The spirit of the decade is a spirited search for environmental quality. To support and sustain this third American revolution will require an educational program as massive as the problem of human survival (15).

REFERENCES

1. Archbald, David, "The Decade in Which Man's Die is Cast," *Wisconsin State Journal*, Madison, January 4, 1970.
2. Brennan, Matthew J., "Closing Summary by the Chair-

man," *Transactions of the Thirty-Second North American Wildlife and Natural Resources Conference*, Wildlife Management Institute, Washington, D. C., 1967, p. 3.

3. "Conservation Education—An Imperative Need," Natural Resources Council of America, Committee on Conservation, Washington, D. C., 1967, p. 6.
4. Cooley, Richard A., *Graduate Education and Research in Natural Resources Public Policy at The University of Washington*, The University of Washington, Seattle, 1966, p. 4.
5. Gardner, John W., "Universities as Designers of the Future," *Educational Record*, Fall 1967, p. 318.
6. Nixon, Richard, quoted in *Time*, February 2, 1970, p. 7. (Special issue on "The Emerging Science of Survival.")
7. Pinchot, Gifford, *The Fight for Conservation*, Harcourt Brace, Garden City, New York, 1910, p. 42.
8. *Report to the President and the President's Response*, White House Conference on Natural Beauty, U. S. Government Printing Office, Washington, D. C., May 1965, p. 42.
9. Roth, Robert, "Fundamental Concepts for Environmental Management Education," *Environmental Education*, Spring 1970, p. 65.
10. Schoenfeld, Clay, *Course Concept*, The University of Wisconsin, Conservation Education Programs, Madison, Wisconsin, 1967, p. 1.
11. Schoenfeld, Clay, "The Ecology of the New Conservation," *Transactions of the 35th North American Wildlife and Natural Resources Conference*, Wildlife Management Institute, Washington, D. C., 1970, in press.
12. Schoenfeld, Clay, "Educating the Public in Natural Resources," *Journal of Soil and Water Conservation*, 23:(No. 6)224, November–December 1968.
13. Schoenfeld, Clay, "Environmental Education and the University," *Educational Record*, Summer 1968, Washington, D. C., p. 305.
14. Schoenfeld, Clay, "Regional Environmental Education Centers," *Environmental Education*, Winter 1970, p. 4.
15. Schoenfeld, Clay, "The Third American Revolution," The University of Wisconsin, Conservation Education Programs, Madison, Wisconsin, 1967.
16. Schoenfeld, Clay, "The University's Changing Role in Conservation," *Transactions of the Thirty-Second North American Wildlife and Natural Resources Conference*, Wildlife Management Institute, Washington, D. C., 1967, p. 179.
17. Schoenfeld, Clay, "What's New About Environmental Education?" *Environmental Education*, Fall 1969, p. 7.
18. Stapp, William B., "Consideration of a Few Approaches to Environmental Education," The University of Michigan, School of Natural Resources, Department of Conservation, Ann Arbor, Michigan, June 18, 1967, p. 3.

ENVIRONMENTAL EDUCATION

A New Journal—A Scholarly Approach to the Study of Man and His Environment

Environmental Education is a quarterly designed to build public interest, understanding, and support in the quest for environmental quality for mankind. It is for people systematically studying the new interpretation of our resources and realizing that pervasive degradation of the environment demands action. The 48-page journal publishes research articles, project reports, and critical essays designed to advance the scientific study of conservation communication and improve field practice in environmental education.

COMING ISSUES

The Universities Encompass Environmental Studies
The Schools Redefine Conservation Education

Subscription Rates:
One year $7.50, student $5.00
Two years $12.00, single copy $2.00

For more information, write:
Environmental Education
Dembar, Box 1605-JR
Madison, Wisconsin 53701

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Levels of Cheating Under Conditions of Informative Appeal to Honesty, Public Affirmation of Value, and Threats of Punishment

CONSTANCE T. FISCHER,[1] University of Kentucky

ABSTRACT

As a means of delineating conditions conducive to minimal classroom cheating, 135 boys and girls in the fourth, fifth, and sixth grades were given an opportunity to cheat on a "General Achievement Test" under Control, Informative Appeal to Honesty, Public Affirmation of Value (not cheating), and two Threat of Punishment conditions. A majority cheated in the first two conditions, but a significant decrease from the Informative Appeal level occurred in all the last three conditions; and among these three, there was no difference. In view of other studies and developmental theory, Public Affirmation of belief in not cheating is posited as the preferable classroom method of minimizing cheating. Methodological procedures for follow-up studies are proposed, with special emphasis on the necessity of ascertaining the experimental variables' personal meanings to the subjects.

CHEATING on school examinations results in inaccurate evaluation of achievement as well as an attenuation of effort toward mastery of the subject matter on the part of the student. Moreover, the occurrence of cheating indicates a failure in the teaching of moral conduct. Beyond these problems, educators are becoming still more concerned that they themselves may contribute to lowered moral standards by allowing the student to discover the efficacy of cheating in our school system (13, 5). These fears have been confirmed in part by Mills' (10) research indicating that sixth-grade children who yield to the temptation to cheat are afterward more lenient toward cheating. It would seem imperative, then, that if only for the sake of future character, present cheating should be obviated.

However, as Kohlberg (8) points out in his review chapter on development of morality, there are few empirical, pragmatic leads to follow toward such a goal. Current research has centered instead on the controversy about generality versus specificity of honesty, on personality characteristics of the cheater, on relevant child-rearing practices and parental traits, class and culture variables, sequential development of moral judgment, timing of punishment, the role of reinforcement variables, and measurement of post-transgression remorse and guilt. The few practicable variables that have been clearly described must be stated in negative terms: physical punishment, character education classes, and moralizing are ineffective in reducing dishonesty. The only variables clearly identified as effective are decreased motivation and decreased opportunity to cheat (2, 6). The only other variable of ready pragmatic relevance has been that of honor codes; demographic reports generally contend that where such a system is working well, empirical levels of cheating are lower than those reported in various other studies.

The purpose of the present study, then, was to determine whether five variations in practical classroom conditions would result in different levels of cheating. An attempt was made to vary these conditions along the dimension of arousal of a personally-held value of not-cheating. The goal was thus to contribute an initial, controlled study which would specifically compare effectiveness of various degrees of positive involvement on the part of middle-age children. Previous experiments have touched on received praise as a variable, but the primary emphasis has been on various forms of threatened and actual punishment.

Method

Subjects in the major part of the experiment were 135 boys and girls enrolled in the fourth, fifth, and sixth grades of three representative public schools in the Lexington, Kentucky area. They were said by their teachers to be of low-to-high middle socioeconomic class.

Cheating was defined as the subject's use of an answer page rather than his own knowledge while taking a 60-item multiple choice "General Achievement Test." Half of the test items were constructed so that either none or all of the five choices was correct, but their arbitrarily designated answers appeared on the back of the test booklet along with the answers to the thirty legitimate questions. Examples of the sham questions are:

(1) *The pepper grinder was invented by: (A) Edison; (B) Franklin; (C) Bell; (D) Faragut; (E) Marconi.*

(2) *Feranti is really a: (A) tree; (B) flower; (C) weed; (D) shrub; (E) grass.*

Since there was no factually accurate answer to these sham items, that designated on the answer sheet could be arrived at only by guessing or cheating. Although several other forms of deception occurred, the criterion of cheating was the occurrence of nine or more sham items answered "correctly." The binomial probability of this number of "correct" sham answers occurring by chance alone is .06; hence, nine or more "correct" sham items is assumed to reflect cheating rather than guessing. During follow-up work, only three of ninety-eight subjects given the test without the answer sheet scored more than eight "correct" sham answers.

Groups of ten to fifteen students were tested at a time, with the individual members always widely separated from each other. This separation was intended to maximize felt safety of cheating, while ostensibly serving to minimize cheating by copying. However, later experimentation indicated that physical separation has no effect on the number of students who cheat. The (female) experimenter was presented as a student teacher helping the regular teachers with the testing. Since pilot data showed that only negligible cheating occurred in the absence of motivation such as the following, subjects were told that their scores would contribute to their 6-weeks grades in social studies. As a further effort to provide awareness both of the possibility of cheating, and of what constitutes cheating, subjects were told that they had 30 minutes to take the test, and that when finished they were to score it; they were told not to look at the answer page before they were ready to score, and not to change answers as they scored since these behaviors constitute cheating.

Control Condition. The Control groups received only the above instructions.

Informative Appeal to Honesty Condition. In addition to the above instructions, these groups were told the importance of obtaining true measures of their knowledge. The explanation was given in terms of the teacher's need to evaluate the effectiveness of her teaching techniques and materials in order to improve them for next year's students. An appeal was also made to the present students' need for evaluation and further teaching help. The tone of the explanation was that of matter-of-fact appeal and information-giving, with the apparent expectation of cooperation.

Public Affirmation of Value Condition. While ostensibly waiting for the proper testing time to begin, the examiner casually asked what grammar school students think about cheating. After general discussion among the students, the examiner asked each in turn to give his name and reasons why he did not think he would cheat on a test like the one he was about to take.

Value-relevant Threat of Punishment Condition. The additional instructions for these groups were:

> Children in your grade have tended to cheat on this test. They get answers to questions they don't know from the answer page. If I should find that any of you has cheated, he or she will have, as a punishment, to write a sentence 50 times. The sentence would be, "Although *I* do not believe in cheating, *I* cheated on this test."

Non-value-relevant Threat of Punishment Condition. Here, the additional instructions were:

> Children in your grade have tended to cheat on this test. They get answers to the questions they don't know from the answer page. If I should find that any of you has cheated, he or she will have as a punishment, to write numbers from one to 100, 25 times.

Earlier paired-comparison ratings by thirty fifth graders of eight possible punishments yielded no significant difference in the judged degree of severity of these two types of punishment.

Semantic Differential Group. Ninety-three fourth, fifth, and sixth grade students in an additional school (not part of the main experiment) completed a previously standardized semantic differential booklet prior to participating in a Non-value-relevant Threat of Punishment condition. The booklet contained fifteen items to be rated on the five (7-point) scales having the highest factor loadings on the evaluative dimension and low loadings on the activity and potency dimensions (11). The five items used to determine moral attitude scores were: myself, cheating on a test, copying someone else's homework, keeping money you find, and not being fair in a game.

Following completion of test-taking in each school, all subjects were brought together to be given an explanation of the test. Group discussion followed. In two of the schools, before and again after the explanation–discussion period, subjects were asked to write "yes" or "no" on a piece of paper to acknowledge anonymously whether or not they had cheated on the test.

Results

Fifteen of twenty-three subjects (65%) cheated under the Control condition (see Table 1). Similar data were obtained in later Control groups tested in another school system by a different examiner. The proportion of subjects who cheated under the Informative Appeal condition did not differ significantly. The proportions of cheaters in the Public Affirmation of Value, Value-relevant Threat of Punishment, and Non-value-relevant Threat of Punishment conditions were all significantly lower than in the Informative Appeal condition ($\chi^2 = 7.22$, $p < .01$; $\chi^2 = 7.68$, $p < .01$; $\chi^2 = 4.71$, $p < .05$), but did not differ from each other. Again, highly similar proportions of cheaters have been found in later groups tested under slight procedural changes (12). Proportions were also highly similar for the respective conditions within each school.

Although subjects appeared able to respond discriminatingly to the semantic differential booklet, there was no significant effect in an analysis of variance of three levels of frequency of cheating. The unit of analysis for each subject was the mean scale value of the difference between the four semantic differential concepts constituting an unethical behavior, and the concept, "myself."

Prior to being told the true purpose of the testing, of the total eighty-six cheaters in the two relevant schools, only six subjects anonymously acknowledged that they had cheated. Following the explanation, and a statement of how many (about half) of the group present had cheated, only twenty-five subjects anonymously acknowledged having cheated.

There were no signficant differences between males and females for either number of cheaters or for the total number of cheating responses. This was true both between conditions and overall. Nor were there differences, either across or within conditions, in the proportions of cheaters or amount of cheating among the three grades.

There was no difference between IQ's of cheaters and noncheaters. However, in one school a (nonsignificant) product–moment correlation of $-.29$ ($N = 43$) was found between number of cheating responses and Otis Form B IQ scores. In a second school, among those cheaters for whom Kuhlman–Anderson Form D IQ scores were available ($N = 21$), the above trend was statistically significant ($r = -.48$, $p < .05$).

Discussion

Given the relatively optimal motivation and opportunity provided by the experiment, nearly two-thirds of the Control subjects cheated; this level of cheating was not significantly reduced by Informative Appeal to Honesty. Both of these findings are consonant with previous frequency data and with numerous observations of a developmental as well as a persistent lag between cognitive understanding and pursuant behavior. In the present study, such a difference was directly evident in the lack of significant relationship between attitude toward unethical behavior (semantic differential ratings) and conduct under a Threat of Punishment condition.

That Threats of Punishment resulted in significantly fewer students cheating, as compared to the Informative Appeal groups, is also consonant with earlier reports and with developmental understanding of children. The finding of direct, constructive importance, however, was that Public Affirmation of the Value of Not Cheating also resulted in a significant reduction in the num-

Table 1.—Proportions of Students Who Created

		Conditions				
Sex	Number of subjects, and % cheaters	Control	Informative Appeal	Public Affirmation	Value-relevant Threat	Non-value-relevant Threat
Boys	N	16	12	16	13	69
	%	56.2	83.3	43.7	30.8	50.7
Girls	N	7	13	15	17	66
	%	85.7	69.2	26.7	52.9	48.4
Over-all	N	23	25	31	30	135
	%	65.2	76.0	35.4	43.3	49.6

ber of cheaters. While the latter condition was no more effective in this regard than the Threat of Punishment conditions, it is *as* effective, and hence on other grounds can be posited as the preferred method for minimizing classroom cheating.

Specifically, a public affirmation situation encourages the child to adopt the value as his own, to commit himself in action as well as belief. In contrast, threats of punishment probably foster an understanding of honesty as an external regulation policed by others. In this latter circumstance, it seems likely that not getting caught is more salient to the student than not cheating. Indeed, there is ample evidence (e.g., 4, 14) that children tend not to identify with parents who inflict physical punishment, and that parents of delinquents use physical punishment more often than do parents of nondelinquents (4). In addition, there is evidence that teachers who are perceived as punitive foster lenient personal attitudes toward misconduct (9). At any rate, to remain effective, threats of punishment would require the backing of an actual punishment system, with its attendant complications for the teacher.

Finally, it may be noted that public affirmation of personal belief in the value of not cheating approximates participation in an honor system—which, when functioning well, is said to minimize cheating, at least among older students. Nevertheless, there is a need for longitudinal studies of the long-term effectiveness of honor systems as well as of the present Public Affirmation approach. In addition, since the literature indicates few instances of changed level of morality during the middle years of childhood, a downward extension of the experimental conditions is also called for.

Post-experimental discussions indicated that subjects were keenly aware of operational definitions of cheating, even though their understanding of ethics was understandably childish. Especially since so few students were able or willing to admit anonymously to having cheated, it seems all the more important that Jones (7) and Kohlberg (8) be followed in their independent recommendations that concrete ethical concepts frequently be discussed and put into specific practice in the classroom. Presumably, this would not only sharpen a cognitive grasp of ethics, but more importantly it would provide repeated opportunities for the child to integrate his beliefs and behaviors. Furthermore, this practice would allow the student to prehend his teacher as valuing honest behavior as well as academic grades.

In summary, it has been the premise of this study that a lived history of honest behavior probably precedes development of an active commitment to such a way of life. Hence, considering that threats of punishment and public affirmation of belief in honesty equally approach this desired level of minimal cheating, then the latter seems more conducive to a personally actualized belief in honesty. In turn, personally held, actualized belief seems more desirable than transient conformity to externally imposed standards.

However, while lower levels of cheating clearly were attained in the Public Affirmation condition, and while the phrase "Public Affirmation" is an accurate operational description, the critical components of the condition are not explicit. That is, as Bennett (1) has postulated, the effective elements of public commitment may be the combination of decision-making and perception of group concensus.

The more complex design of the Threat of Punishment conditions renders their critical components still less certain. The facts are that both conditions led to significantly less cheating than in the Informative Appeal condition, and that there was no difference between the Threat conditions nor between them and the Public Affirmation condition. But we cannot know whether subjects actually prehended one Threat condition as value-relevant and the other as not. In retrospect, it seems that the implied possibility of being "found out" may have aroused personal values in both conditions. Also, it could be that fear of punishment itself superseded the intended difference in degree of value-relevance.

The incidental data showing no sex or grade differences in frequency of cheating are consistent with the majority of other studies, and they are equally inconclusive. That is, while the same test was given to all students, it was then necessarily more difficult for the younger subjects—who are sometimes found to cheat less than the higher grades, unless under increased motivation. With reference to absence of a sex difference, it should be noted that under other circumstances where the meaning of success on a test might differ for boys and girls, their relative level of cheating might also vary.

Similarly, in the absence of an understanding of just what the test situation meant to each student, the generality of the IQ findings is limited. Indeed, without such an understanding, it will continue to be difficult to establish firm trends among large samples, and where they are established, their meaning and practical significance are severely attenuated. Within this context, then, the present data are that given the experimental conditions, there was no significant relationship between IQ and decision to cheat; but among cheaters, lower IQ seemed to be associated with a higher degree of cheating. The former finding is consistent with a minority of published reports (the majority show a significant if small positive correlation between honesty and IQ), while the latter finding has not received previous attention.

Limitations

The present study provides both a solid ground for further research and specific, immediately practicable implications for classroom application.

Nevertheless, two major limitations should be spelled out. First, in view of Mills' (10) finding that students who yield to the temptation to cheat are afterward more lenient toward cheating, the ethical status of the present experiment in retrospect is itself questionable. Perhaps this kind of attempt to study the phenomenon has resulted in less critical attitudes toward self and cheating, which may in turn be conducive to future cheating. Furthermore, since students tend not to admit their behavior, constructive use of post-experiment discussion and confrontation is limited. Natural field studies of on-going behavior may be the only ethical procedure.

Second, as in all other relevant published studies, the personal meanings to the individual subjects of both the experimental conditions and of cheating itself were not directly studied. Following the present study, Fischer (3) demonstrated elsewhere the necessity and feasibility of determining the personal meaning-context within which measured variables occur. With subjects in the middle years of childhood it is recommended that taped individual interviews be conducted, encouraging each subject to describe in his own way both cheating as such and his experimental condition as he prehended them. Only with such data will the critical components of the various conditions emerge directly. Indeed, such data is probably the key to synthesizing the generality-specificity controversy over the nature of honesty. That is, if various situations have common meanings to a subject, he will probably behave in such a way that cheating is seen as a general trait. In those situations whose personal meanings differ for him, however, his behavior would probably reflect honesty as situation-specific.

FOOTNOTE

1. Now at Duquesne University.

REFERENCES

1. Bennett, E. B., "Discussion, Decision, Commitment, and Concensus in 'Group Decision,'" *Human Relations*, 8:251–274, 1955.
2. Drake, C. A., "Why Students Cheat," *Journal of Higher Education*, 12:418–420, 1941.
3. Fischer, C. T., "Male-Female Differences in Catholic College Students' Attitudes toward the Possibility of Social Change," *The Catholic Psychological Record*, 6:30–34, 1968.
4. Glueck, S.; Glueck, E., *Unraveling Juvenile Delinquency*, Harvard University Press, Cambridge, Massachusetts, 1950.
5. Harris, E.; Harris, D., "How We Encourage Cheating; Symposium, with Study-Discussion Program," *PTA Magazine*, 60:35–36, February 1965.
6. Howells, T. H., "Factors Influencing Honesty," *Journal of Social Psychology*, 9:97–102, 1938.
7. Jones, V., "Character Development in Children: An Objective Approach," in Carmichael, L. (ed.), *Manual of Child Psychology*, John Wiley and Sons, New York, pp. 707–751, 1954.
8. Kohlberg, L., "Development of Moral Character and Moral Ideology," in Hoffman, M. L.; Hoffman, L. W. (eds.), *Review of Child Development Research*, Russell Sage Foundation, New York, 1964, Vol. I, pp. 383–431.
9. Kowin, J. S.; Gump, P. V., "The Comparative Influence of Punitive and Nonpunitive Teachers upon Children's Concepts of School Misconduct," *Journal of Educational Psychology*, 52:44–49, 1961.
10. Mills, J. R., "Changes in Moral Attitude following Temptation," *Journal of Personality*, 26:517–531, 1958.
11. Osgood, C. E.; Suci, G. J.; Tannenbaum, P. H., *The Measurement of Meaning*, University of Illinois Press, Urbana, 1957.
12. Taylor, C. I., "An Experimental Enquiry into Test-Cheating Behavior," master's thesis, University of Kentucky, 1963.
13. Weldon, L. L., "Cheating in School; Teachers are Partners in Crime," *Clearing House*, 40:462–463, 1966.
14. Whiting, W. M.; Child, I., *Child Training and Personality: A Cross-Cultural Study*, Yale University Press, New Haven, Connecticut, 1953.

Forthcoming Articles

Dahlem, Glenn G.; St. Louis Missouri
"ACTUAL. IDEAL, AND EXPECTED ROLE CONCEPTS OF SECONDARY COUNSELORS"

Garfield, S. Jeffrey, Helen A. Cohen, Robert M. Roth, Harris L. Berenbaum; Illinois Institute of Technology, Chicago.
"THE EFFECTS OF GROUP COUNSELING ON CREATIVITY"

Glick, Oren; Seattle University, Poulsho, Washington
"THE INTERACTION EFFECTS OF SEX OF MOTHER'S CHILD AND CHILD'S READING PERFORMANCE ON THE MOTHER'S EVALUATION OF THE SCHOOL"

Golub, Lester S.; University of Wisconsin, Madison
"A MODEL FOR TEACHING COMPOSITION"

Loadman, William E.; Michigan State University, East Lansing
"THE EFFECTS OF A MEDIATIONAL DEVICE ON PAIRED-ASSOCIATE LEARNING; LEARNING AND RETENTION OF SIXTH GRADE STUDENTS"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Credibility of Information for Educational Innovation

FRANK B. MURRAY,[1] University of Delaware

ABSTRACT

An aspect of the apparent independence of educational innovation and educational research was investigated. Thirty-four professors of educational psychology and fifty junior high school teachers ranked in importance eight kinds of information about an educational innovation and evaluated the credibility of eight sources of educational information. The correlation between the rankings by each faculty was insignificant, but the correlation between the credibility evaluation by each faculty was .84 and significant. Each faculty generally found its own colleagues and their activities more important and credible; an important finding was that teachers regarded standardized test research results as much less important than the professors who regarded it as most important.

LEWIS Terman (4) lamented that after nearly 50 years of educational research to the contrary, educators still refused to accelerate bright children because of the alleged deleterious effects of acceleration on social adjustment. An impressive array of educational research on the effects of school attendance, television, independent study, correspondence courses, class size, counseling, school size, teacher training, ability grouping, etc., on school learning has been assembled by Stephens (3), and the findings of this research are remarkably unrelated to educational practice. Gillooly's (1) critical review of the research on i.t.a. indicated that the adoption of that innovation in reading instruction would have to be based on criteria other than good research.

One aspect of the apparent independence of educational innovation and educational research may be that educators and researchers are sensitive to different kinds of information. For whatever reasons, the information that one group thinks is critical, the other apparently thinks is irrelevant or unreliable. In the present investigation, members of a university research faculty in education and a junior high school faculty were asked to rank eight kinds of information in terms of how important each would be in their decision to adopt a particular educational innovation. A second part of the investigation required the same subjects to evaluate some standard sources of information about educational matters.

Method

Subjects. Thirty-four faculty members of the Department of Educational Psychology of the University of Minnesota (25 PhD, 6 MA, and 3 BA), and fifty faculty members of the Burnsville, Minnesota Junior High School (10 MA, 29 BA, 3 AA, and 8 high school diploma) completed a questionnaire that required them to rank eight kinds of information and to judge which of two sources of educational information was most credible. The mean number of years the university faculty had taught was 9.36 years, and the high school faculty, 4.42 years.

Questionnaire. The instructions for the first part of the questionnaire were: *If you had to make a decision about whether to adopt a certain educational innovation (such as team teaching, computerized instruction, new math, initial teaching alphabet), you would no doubt want to know certain things about the innovation.*

On the next page there are several kinds of information you might use. Please rank these according to how important or valuable they would be in your decision. Put a "1" next to the piece of information that is most important, a "2" next to the one that is next in importance, etc. On the next page were listed one under the other the following kinds of information:

1. *The results of a questionnaire given to school administrators (Principals, Supervisors, Superintendents, etc.) who have supervised the use of the innovation. The information could be in the form of a percentage of administrators who endorsed the innovation.*
2. *Your own observation of the practice or use of the innovation in a classroom.*
3. *A comparison of students school grades before and after the innovation, or a comparison of school grades between a group subjected to the innovation and an equivalent*

group of students using the older method or practice.

4. *Logical considerations about the innovation, such as the reasonableness of the arguments put forth to justify its adoption; the rationale or philosophy of the innovation, or the reasons why it is supposed to work; the theories from which the innovation is derived.*
5. *The results of a questionnaire given to teachers who had actually used or practiced the innovation. This information could be in the form of a percentage of teachers who have tried the innovation, believe it, prefer to continue using it, and think it is superior to former practices.*
6. *The endorsement of the innovation by well-known and highly respected educators.*
7. *The personal endorsement of the innovation by colleagues who are good friends of yours and who would give you their personal opinion of the innovation.*
8. *A comparison of standardized tests results between equivalent groups of children who have and have not been subjected to the innovation, or a comparison of standardized test results for a group before and after the innovation.*

In the second part of the questionnaire, subjects were instructed: *Suppose that you are in an administrative position which requires you to make a decision about an educational issue that you know very little about. When you turn to some standard sources of information, you find they disagree, and you have to choose between them. Circle the one you would choose to believe.*

For example: (a truth-teller) and a liar
In all of the decisions you make on the following page, you should assume that each item is one of the best in its category.

On the next page the following eight sources of information were presented in all possible pairs (28): Educational Research Journal, Professional Teachers' Journal, College Textbook, T.V. Special, Weekly Newsmagazine, College Professor, U. S. Office of Education Bulletin, and Newspaper.

Results

The mean rank and standard deviation of each kind of information is presented in Table 1. Junior high school faculty ranked the results of a questionnaire given to teachers and the endorsement by well-known educators significantly higher in importance than did the university faculty ($p < .01$ and $p < .05$ respectively). The university faculty ranked the results of the standardized test comparison much higher in importance than did the junior high school faculty ($p < .001$).

The order of importance (high to low) of the eight items in the decision of the university faculty were: (1) standardized tests; (2) logical reasons; (3) own observation; (4) teacher questionnaire; (5) school grades; (6) good friend; (7) educator's endorsement; and (8) principal questionnaire. For the junior high school faculty the order was: (1) teacher questionnaire; (2) own observation; (3) logical reasons; (4) educator's endorsement; (5) school grades; (6) standardized test; (7) good friend; and (8) principal questionnaire. The Spearman rank order correlation between the two rankings was .45 which is not significantly different from a zero correlation ($p > .05$). The mean product moment correlation of each subject's ranking of the items with all other subjects' rankings was .14 for the junior high faculty and .54 for the university faculty.

Analysis of the twenty-eight comparisons of the eight sources of information indicated that the order of credibility of the items for the university faculty (from most credible to least) was:

Table 1.—Mean Rank and Standard Deviation of Eight Kinds of Information for University (N=34) and Junior High Faculty (N=50)

	Kinds of Information							
	Principal Questionnaire		Personal Observation		School Grades		Logical Reasons	
Faculty	Mean	SD	Mean	SD	Mean	SD	Mean	SD
University	6.06	2.09	4.02	2.18	4.65	2.48	3.62	2.10
Junior High	5.85	2.01	3.36	2.22	4.83	2.19	3.83	2.47

	Teacher Questionnaire [1]		Educator Endorsement [2]		Good Friend		Standardized Test [3]	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
University	4.41	1.69	5.56	1.81	5.06	1.92	2.61	2.24
Junior High	3.06	1.93	4.72	1.66	5.42	2.01	4.91	2.33

Table 2.—Numbers of University (N=34) and Junior High Faculty (N=50) That Chose to Believe Each Source of Information in Selected Conflicting Pairs

	Paired Sources of Information									
	Mag. vs. Prof.		Teach. J. vs. Prof.		T.V. vs. Prof.		News vs. Prof.		U. S. vs. Prof.	
Faculty										
University	3	31	10	24	2	32	0	34	17	17
Junior High	25	25	39	11	20	30	12	38	39	11
	Mag. vs. Text		Teach. J. vs. Text		T.V. vs. Text		News vs. Text		U. S. vs. Text	
University	1	33	15	19	3	31	1	33	18	15
Junior High	15	35	38	12	19	31	12	38	42	8

(1) educational research journal, (2) college professor, (3) U. S. Office Bulletin, (4) college text, (5) professional teachers' journal, (6) T.V. special, (7) newsmagazine, and (8) newspaper. The credibility order for the junior high school faculty was largely the same with exception that the college professor placed fifth, not second, and the professional teachers' journal placed third, not fifth. The Spearman rank order correlation between these two orders was .84 which is a significant correlation ($p < .01$).

An analysis of the differences in the proportions of university and junior high school faculty that chose to believe one or the other member of each pair indicated that there was significant disagreement on eleven of the twenty-eight items. Five of these were disagreements about the credibility of the college professor (Table 2) when he was compared to the newsmagazine ($\chi^2 = 13.6$, $p < .001$), the professional teachers' journal ($\chi^2 = 17.7$, $p < .001$), the television special ($\chi^2 = 10.5$, $p < .001$), the newspaper ($\chi^2 = 7.6$, $p < .01$), and the U. S. Office of Education Bulletin ($\chi^2 = 5.9$, $p < .05$). In each case the university faculty more than the junior high faculty tended to believe the college professor. Five other significant disagreements (Table 2) between the university and junior high faculties were about the credibility of the college text when it was compared to the newsmagazine ($\chi^2 = 7.95$, $p < .01$), professional teachers' journal ($\chi^2 = 7.51$, $p < .01$), the T.V. special ($\chi^2 = 7.46$, $p < .01$), the newspaper ($\chi^2 = 5.35$, $p < .05$), and the U. S. Office of Education Bulletin ($\chi^2 = 7.20$, $p < .01$). In each case, the university faculty valued the textbook more than the junior high faculty did. The final disagreement between the faculties was over the credibility U. S. Office Bulletin and the educational research journal ($\chi^2 = 6.59$, $p < .05$) in which the journal was chosen by thirty-one of the university and thirty-two of the high school faculty and the Bulletin by eighteen of the junior high and only three of the university faculty.

Discussion

In a real sense, the subjects indicated only what they thought were the most important kinds of information and most credible sources of information, and not how they would actually use them if they made the decision to adopt an innovation. Since what people think they ought to do in some way places a ceiling on what they may in fact do, the results of this study are all the more interesting.

Most interesting is the large disagreement about the importance of standardized test results between the educators and the researchers. This disagreement is perhaps symptomatic of a critical difference in cognitive style between the researcher and practitioner. The researcher has chosen, if nothing else, to be certain of what he is speaking about; to do this he operationalizes his constructs. Often the operationalized definition, such as a standardized test, is so far removed from the rich theoretical and connotative meaning of a construct that the educator cherishes that he often dismisses the best educational and behavioral research as somewhat interesting, but irrelevant. For example, the educator is simply not willing to say, as the researcher must, that, intelligence is what the Stanford–Binet, or some other set of tasks, measures. In the daily crisis of teaching and governing the school, the educator cannot afford the luxury of being certain, and so he is forced quite literally not to know exactly what he is talking about.

The disagreement between the two faculties over the importance of the teacher questionnaire and the endorsement of well-known and respected educators is in accord with the interpretation given for the standardized test differences. It perhaps is to be expected that the teachers would value the opinion of their colleagues more than the professors who, as the results of the second part of questionnaire indicated, also place greater value on the opinions of their colleagues means of a combination of mental and social maturity.

The intercorrelations among academic achievement, mental age, and social age were computed and are presented in Table 1. From the data it may be seen that there exists a high relationship between mental age and academic achievement, while the correlations between social maturity and each of the variables of mental maturity and academic achievement may be described as moderate.

Table 1.—Intercorrelations Among Measures of Academic Age, Mental Age, and Social Age

Measure	X_1	X_2	X_3
X_1 Academic Age	1.000	.791*	.344*
X_2 Mental Age		1.000	.403*
X_3 Social Age			1.000

*Significant at the .05 level.

In view of the relationship found to exist between mental and social age, it must be concluded that common elements are functioning in the determination of both mental age and social age. This conclusion is warranted by the results of the computation of the partial correlation between academic age and mental age with social age held constant. The drop from .791 to .759 in the coefficient of correlation between academic age and mental age reveals the relatively small, although measurable, influence exerted by social age.

It was further found that while mental age accounted for approximately 55 percent of the total variance in academic achievement scores, the inclusion of social age in the regression equation increased the predictive value of mental age by only .15 percent. Therefore, it was concluded that the social maturity of normal children, as measured by the Vineland Social Maturity Scale, was of no significance in predicting academic success.

The search for significant predictors of academic achievement will undoubtedly continue, especially considering the present state of our educational system and the desire of educators to provide for the individual needs and problems of all students. Suitable educational objectives and teaching procedures can be more readily and accurately established when the probable achievement level of pupils is known. Therefore, still other factors, including other types of maturity, must continue to be investigated in order to ultimately resolve the problem of prediction.

FOOTNOTE

1. This report is part of a larger study conducted by the author while an NDEA Fellow at Ohio University. The original report, entitled "The Role of Social Maturity in the Prediction of Academic Achievement," is on file in the Education Library, College of Education, Ohio University, Columbus.

REFERENCE

1. Doll, Edgar A., *Measurement of Social Competence*, Educational Test Bureau, Minneapolis, Minnesota, 1953, p. 2.

ERRATA

The "Summary and Review of Investigations Relating to Reading," July 1, 1968 to June 30, 1969, by Harris and others in *The Journal of Educational Research* (March 1970), p. 301, Column 1, lines 10–15 should read:

The effect of measures of visual–factual–kinesthetic integration to differentiate adequate from inadequate readers was explored by Hurley (3). Forty pairs of pupils in grades 2 and 3, one an adequate reader and one an inadequate reader, were matched according to sex, chronological age, and IQ score. Twenty-four pairs of pupils were selected from one community, sixteen pairs from another.

The reviewer is grateful to the author for calling his attention to this matter.

T. L. Harris

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Factors Relating to the Innovativeness of School Systems[1]

LEO R. HILFIKER
The University of Wisconsin, Madison

ABSTRACT

System innovativeness was determined through the use of rankings for each district from three distinct sources: the district superintendent, the professional staff of the system, and a ranking derived from a panel of ten experts. Selected sections of an instrument designed by the Cooperative Project in Educational Development (COPED) were used to obtain data regarding the interpersonal behavior (independent) variables. Spearman's rank order correlation was used for determining the relationships between the dependent and the independent variables. Independent variables having a significant relationship ($p < .05$) to school system innovativeness were: social support provided by the principal as perceived by the professional personnel (.01), the perceived problem-solving adequacy of staff meetings (.01), satisfaction with the amount of time devoted to problem-solving in staff meetings (.05), perceived powerlessness in system faculty and administrative council meetings combined as a single variable (.01), and openness (.05) and trust (.01) as interpersonal process norms of the system as perceived by professional personnel.

IN THIS STUDY, the investigator explored some of the variables associated with the human environment in educational organizations in terms of how these variables might be related to the innovativeness of school systems. In the discussion of the study, certain terms are used which are not always found in educational literature. Some of these terms are defined here in order to clarify their meaning.

Interpersonal Process Norms—implicit standards of interpersonal behavior perceived as expected or agreed upon by members of a group as a result of their mutual expectations of each other.

Problem Solving Adequacy—the degree to which meetings are characterized by the clarity and control of the meeting agenda, the diagnosis and definition of problems, the generation and discussion of possible solutions, the resolution of problems through decision-making, and the implementation and evaluation of action steps.

Executive Professional Leadership (EPL)—the degree to which teachers perceive the principal as stressing his obligation to improve the quality of staff performance. (Adapted from Gross and Herriott.)

Social Support—the degree to which teachers perceive the principal as a warm, socially responsive individual who tends to create an empathic and non-threatening environment. (Adapted from Gross and Herriott.)

Powerlessnes—a quality or state of being devoid of strength, authority, or resources to act or influence others.

Openness—a quality or state of being characterized by ready accessibility, cooperative attitudes, tolerance of internal change, and permissiveness of diversity in social situations.

Trust—the degree to which an individual perceives interpersonal relationships as characterized by an assured reliance or confident dependence upon the character, ability, or truthfulness of others.

Adaptiveness—the degree to which an individual perceives interpersonal relationships as characterized by a ready capability for modification or changes in social conditions, ways, or environments.

Conceptual Framework. The development of an adequate theory linking interpersonal relationship variables to organizational innovativeness is lacking. However, Carl Rogers (1, 3) presented a theory which has considerable relevance to the interpersonal relationship variables and concepts used in the present study. Rogers, in a paper describing conditions which should foster creativity or innovativeness, indicated that conditions of creativity cannot be forced, but must be permitted to emerge. He outlined two general conditions for maximizing the emergence of constructive creativity: (1) psychological safety and (2) psychological freedom.

Psychological Safety. The concept of *psychological safety* is related to *trust.* According to Rogers, the attitude of psychological safety will manifest itself when a "teacher, parent, etc., senses the potentialities of the individual and thus is able to have an unconditional faith in him, no matter what his state. The effect on the individual as he apprehends this attitude, is to sense a climate of safety." Additional ways of providing psychological safety are: genuine empathy, understanding, and an absence of external evaluation. The latter (external evaluation) is often perceived by individuals as a threat and creates a need for defensiveness which, in turn, may result in a diminished degree of trust.

Psychological Freedom. The concept of *psychological freedom* is related to *openness* and to a lesser degree, *powerlessness.* Rogers defines psychological freedom as permissiveness which "gives the individual complete freedom to think, to feel, to be whatever is most inward within himself. It fosters the openness . . . which is a part of creativity." This kind of freedom and openness, according to Rogers, fosters the development of a secure locus of evaluation within oneself. To the degree that an internal locus is possible within a given individual or group, a differential feeling of powerlessness (or power) could develop the explanation being that the further an individual perceives the locus of evaluation removed from himself, the more likely he is to feel powerlessness.

Rogers summarized this theory by stating several hypotheses regarding the fostering of constructive creativity, two of which have relevance for this study:

> 1. Given two matched groups, the one in which the leader establishes a measurably greater degree of conditions (psychological safety and freedom) will spontaneously form a greater number of creative products, and these products will be judged to be more significantly novel.
> 2. A group in which conditions (psychological safety and freedom) are established should, according to our theory, have more effective and harmonious interpersonal relationships than a matched group in which these conditions are present to a lesser degree. (The reasoning is that if creativity is all of a piece, then a group in which the fostering conditions are established should be more constructively creative in social relationships.)

Group and Leadership Relations. A concept expressed in part by Rogers and used in this study is that the interpersonal relationships that develop in a subgroup of the school system (in professional meetings) have an important impact upon the creativity and innovativeness of the members of that group and possibly of the entire system. In the hypotheses presented by Rogers there is no specification or limitation as to the size of the group except whatever practical limitations might arise because of a group's ability to "spontaneously form a greater number of creative products." It would appear that such reasoning could be applicable to several social or working groups throughout a school system, i.e., problem-oriented committees, building faculties, social committees, or the system itself.

Rogers (2) set forth another concept which included leadership as an important factor in the development of a creative group. He indicated that "we know how to establish, in any group, the conditions of leadership which will be followed by personality development in the members of the group. . . . If the leader is acceptant . . . understanding of others in a sensitively empathic way; if he permits and encourages free discussion; if he places responsibility with the group; then there is evidence of personality growth in the members of the group, and the group functions more effectively, with greater creativity and better spirit."

The Hypotheses

Rogers proposed two concepts regarding interpersonal relationships and creativity which have particular relevance for the present study. The first embodies the idea that the climate (conditions) created by a group leader will affect the creative output of the group. The second concept emanates from the first, i.e., if a safe and free climate can be established, a group will not only be more creative but will enjoy more effective and harmonious interpersonal relationships.

These concepts appear to have an analogue in most school situations. The principal normally occupies a leadership position with respect to the building faculty. According to the foregoing discussion by Rogers the creativity of the faculty (group), in terms of the quantity and quality of its innovative product, will be responsive to certain behaviors exhibited by the principal (leader). The behaviors can be used by the principal to establish a psychological climate (conditions) which will enable creativity or innovativeness to emerge and grow. If such a climate can be established then the creativeness that does emerge will not be restricted solely to tangible products but will also result in more constructively creative social relationships among faculty members. The possibility, implicitly raised by Rogers, is that once such a process is begun and nurtured it could become cyclical and self-reinforcing. It is conceivable that the psychological climate or conditions may be initially created or maintained by individuals occupying roles other than that of the principalship or by groups other than building faculty. It may be created deliberately or by chance. The present study focuses upon only one of these possible combinations, the system principals (leaders), the system faculties (groups), and the faculty meetings as a potential locus for the general establishment of psychological conditions thought to be relevant to interpersonal relationships and innovativeness.

Specifically, the questions examined in the present study are stated in the form of null hypotheses and ancillary questions which follow:

1. There is no relationship between school system innovativeness and interpersonal process norms as measured by the degree of adaptiveness, openness, and trust perceived by the professional staff.
2. There is no relationship between school system innovativeness and the executive professional leadership and social support provided by principals as perceived by the professional staff.
3. There is no relationship between school system innovativeness and interpersonal process norms in faculty meetings, as measured by openness and powerlessness as perceived by faculty members.

Ancillary Questions:

a. Is there a relationship between school system innovativeness and interpersonal process norms in administrative council meetings, as measured by openness and powerlessness as perceived by principals?
b. Is there a relationship between school system innovativeness and the degree of satisfaction with the problem-solving adequacy of system professional meetings as perceived by professional personnel?
c. Is there a relationship between school system innovativeness and the degree of satisfaction with the amount of time devoted to problem-solving in pro-

fessional meetings as perceived by professional personnel?

d. Is there a relationship between school system innovativeness and expenditures per pupil?

e. Is there a relationship between school system innovativeness and age of professional personnel.

Design and Methodology

Basic to the study was the thesis that it is possible to determine the degree to which a school system is innovative and that certain interpersonal process norms exist within a given system which are related to system innovativeness. The innovativeness of a school system was determined by utilizing three separate procedures: one based upon a ranking of the innovativeness of the eight systems participating in the study by a panel of ten experts from the Department of Public Instruction; the second, a ranking derived from a quantitative accounting of innovations by professional personnel of a given system with the exception of the system superintendent; and third, a ranking derived from information procured during a comprehensive, structured interview with the school system superintendent. These three rankings were combined into a composite ranking which was used as the data for the dependent variable, school system innovativeness.

The eight school systems and the adult population included in this study are identical to those used in a comprehensive long-term study of planned change by the University of Wisconsin Research and Development Center for Cognitive Learning. The Center, in turn, is informally affiliated with the Cooperative Project in Educational Development (COPED) which published the package of instruments which were administered in the eight school systems. The instruments were extensively revised by the Center project team following the completion of this study.

The sampling procedures were suggested by the national COPED project, i.e., the selected instrument items were administered to a 30 percent sample of the total professional personnel in each of the eight participating school systems. This sample resulted in the following distribution by school systems:

System (Coded)	1	2	3	4	5	6	7	8
Total in Sample	161	59	20	139	34	74	25	111

The COPED Instrument is a collection of questionnaires designed to measure a variety of attitudinal, perceptive, behavioral, and normative social characteristics of professional adults in school systems. It consists of two main parts, each of which is divided into sections. The titles of the various instrument sections are listed below. The column on the right indicates which sections were utilized to supply data for this study.

COPED Instrument, Part I

Instrument Section	*Utilized as a Data Source for*
1. Biographical Information	*Ancillary question (age)*
2. Social Structure	*Not used*
3. School Objectives	*Not used*
4. Degree of Satisfaction	*Not used*
5. School Climate	*Not used*
6. Norms (Do's and Dont's)	*Independent variables (openness, adaptiveness, trust)**
7. Patterns of Relationships	*Not used*
8. Meetings	*Independent variable (openness, powerlessness, problem-solving)**
9. Final Reactions	*Not used*

COPED Instrument, Part II

10. Your Principal	*Independent variable (EPL, social support)*
11. Influence	*Not used*
12. Innovations (by teachers)	*Dependent variable*
13. Reactions to Center Research Team	*Not used*

COPED Instrument (Interview Format)

Superintendent Interview	*Dependent variable*

** Sections 6 and 8 were factor analyzed in order to determine specific variables for use in this study.*

The general methodological procedure used was to compute mean scores for each independent variable from the raw data by school system. The mean scores were used to assign an appropriate rank to the eight school systems on the basis of the scores for each of the independent variables. A rank order correlation was then computed using the rankings of each independent variable taken separately.

The major independent variables investigated were interpersonal process norms perceived to exist in the participating school systems. These norms were analyzed at three main loci: the principal as perceived by the professional staff, the professional teaching staff, and professional staff meetings as a vehicle for school system problem-solving and for the development and integration of collective interpersonal process norms.

In addition to the above, two other independent variables were investigated; the expenditures per pupil and the age of professional personnel in the eight systems. These variables were included in the study when it was found that previous studies which investigated the relationship of school district expenditures and the age of professional personnel to school system innovativeness resulted in inconsistent findings.

The Findings

The findings of this study support some of the hypotheses in varying degrees while other hypotheses are rejected. The null hypotheses were rejected if the level of significance of the rank order correlations was at or beyond the .05 level. A significance level of .10 was regarded as suggestive and the relationship involved was considered worthy of further exploration. The results of the study are summarized here as they related to the order of the hypotheses and ancillary questions.

Hypothesis 1. A significant relationship was found to exist between school system innovativeness and the interpersonal process norms of *openness* and *trust* as perceived by the professional personnel of the system. The portion of hypothesis 1 relating to the norms of *openness* and *trust* was rejected.

No significant relationship was found to exist between the interpersonal process norms of *adaptiveness* and school system *innovativeness.* The portion of hypothesis 1 relating to the norm of *adaptiveness* was accepted.

Hypothesis 2. No significant relationship was found to exist between the *executive professional leadership* of the principals of a school system, as perceived by their professional staffs, and the *innovativeness* of that system. That portion of hypothesis 2, relating to EPL, was accepted. However, the findings indicated a rank correlation of .476, significant at the .10 level, which can be regarded as a suggestive finding and worthy of further exploration.

The *social support* perceived as given to faculty members by principals was found to have a significant relationship to the innovativeness of the school system. The portion of hypothesis 2 relating to *social support* was rejected.

Hypothesis 3. No significant relationship was found to exist between the degree of *openness,* as a perceived interpersonal process norm in building faculty meetings held in a given school system, and the *innovativeness* of that system. That portion of hypothesis 3 relating to the norm of *openness* was accepted.

No significant relationship was found to exist between the degree of *powerlessness,* as a perceived interpersonal process norm in building faculty meetings held in a given school system, and the *innovativeness* of that system. That portion of hypothesis 3 relating to the norm of *powerlessness* was accepted.

Ancillary Question (a): A significant relationship was found to exist between school system *innovativeness* and the interpersonal process norm of *powerlessness* as perceived in administrative council meetings by principals. No significant relationship was found to exist between school system *innovativeness* and the interpersonal process norm of *openness* in administrative council meetings as perceived by principals. Both of the above findings relating to ancillary question (a) should be interpreted with caution because of the small sample used for the variables of *openness* and *powerlessness* in administrative council meetings as perceived by professional personnel.

Ancillary Question (b): A significant relationship was found to exist between school system *innovativeness* and *problem-solving adequacy* of professional meetings as perceived by professional personnel.

Ancillary Question (c): A significant relationship was found to exist between school system *innovativeness* and the degree of satisfaction with the amount of time devoted to *problem-solving* in professional meetings as perceived by professional personnel.

Ancillary Question (d): No significant relationship was found to exist between school system *innovativeness* and *expenditures per pupil.*

Ancillary Question (e): No significant relationship was found to exist between school system *innovativeness* and the *age* of professional personnel.

A significant relationship was found to exist between school system *innovativeness* and the interpersonal process norm of *trust* as perceived by the professional personnel of the system. This finding, and the findings regarding the variables of *openness* and *social support,* appears to support some of the concepts advanced by Carl Rogers and Matthew Miles. The essential argument advanced by Rogers was that if the leader of a group establishes conditions of psychological safety and and freedom, the group will spontaneously form a greater number of creative products which will be more significantly novel and the group will enjoy more effective and harmonious interpersonal relationships. Miles noted that certain phenomena characterize groups which tend to generate change and innovation. He indicated that such groups can be expected to experience high autonomy and spontaneity with freedom for creative experimentation, high quality problem-solving through increased communication, and norms that actively support change. The variable of *adaptiveness* as used in this study, would appear to provide a measure related to the norm of support for change. However, no relationship was found to exist between school system *innovativeness* and the interpersonal process norm of *adaptiveness.* This finding raises a question as to whether individual propensity for change is necessarily reflected at the organization (system) level and if it is not, what factors would account for such a situation.

Implications

Change has become a permanent and integral part of modern society. If education is to become a part of the movement and momentum of social change, then more knowledge is needed regarding the effective ways and means of instituting changes and innovations in school systems. Much of the research dealing with innovation in education has utilized variables connected with the means by which innovations were introduced or diffused. Another area of research emphasis has been directed toward the characteristics of innovators. Little attention has been given to the social or psychological characteristics of the receiving system (such as a school or school system) and how these characteristics might affect the fate of a given innovation or change.

The findings reported in this study suggest that certain interpersonal relationship variables, within the context of organizational climate, may be among the most important variables to consider in initiating and maintaining innovations in educational organizations. The long-term success of school system innovative efforts may be due, to a greater degree than previously suspected, to the social–psychological state of the system's organizational climate. If it becomes possible to consistently diagnose and evaluate the "state" of a school system's organizational climate, it might be feasible to modify the adaptability of professional personnel and to change or create organizational structures and processes which tend to enhance the possibilities of successful institutionalization of innovations. An instrument designed to provide data appropriate to such change processes, with the ultimate objective of modifying the system, might also aid in identifying conditions contributing to excessive change or unstable conditions. An analysis of such conditions might indicate that the system should achieve or return to a state of equilibrium rather than undertake extensive change efforts.

An instrument which a school system could use to assess the existing state of interpersonal relationships, or a change in those relationships over a period of time, might also furnish diagnostic information which would be of intrinsic value to the staff and to the administration. This might be particularly true of interpersonal process norms and problem-solving norms that characterize professional staff meetings. School system self-study and awareness of the degree of openness, trust, etc., perceived to exist throughout the system could be considerable assistance to the personnel of a school system in knowing more about themselves and about the potentialities of the organization. While the cause and effect relationships of the variables used in the present study are not yet clear, the evidence suggests that future research in the area of interpersonal relationships in school systems may be particularly relevant at the present time for two reasons: teachers are beginning to work increasingly within a professional group setting as opposed to that of the isolated self-contained classroom, and teachers are beginning to perceive themselves as a professional group capable of problem-solving and self-improvement at higher levels of decision-making.

The findings of this study indicate a number of other implications for educational practice and future research:

1. Efforts should be undertaken to increase awareness of the importance of human relations skills in classroom as well as organizational settings. Little is known about the importance of interpersonal process norms and interpersonal relationships between teachers and pupils in classrooms. Research is needed which will provide more knowledge about the social–psychological climates of classrooms, how such climates are formed, and how they can be modified.

2. Professional personnel employed by school systems appear to differ in their readiness to accept and support change and innovation. Research is needed which will examine the characteristics of persons who prefer to work in innovative environments and those who prefer to work in stable environments. It is possible that school systems that are either innovative or traditional and unchanging to the extreme could hire personnel who might supply the characteristics which would bring the system into a state of controlled change or equilibrium.

3. Some form of problem-solving training, preferably integrated with human relations skills, appears to be an important consideration in preservice and in-service training programs for educators. The effectiveness of an organization is often dependent upon the ability of the members, individually or collectively, to solve problems. If time spent in professional meetings is perceived by participants as a waste of time or if participants have a sense of "going through the motions" or if they feel powerless to make adjustments in the system, then the problem-solving potentialities of professional meetings may be adversely affected.

4. The advancement of innovative practices might be enhanced through the development of diagnostic instruments which could be administered, scored, and evaluated by personnel in the school system. Such procedures might identify administrators, teaching teams, building faculties, or professional personnel in the organization who would be particularly interested in and supportive of change efforts and innovative practices. The identification of innovative personnel could lead to the formation of a change agent team whose primary role would be to plan and coordinate processes of change within the school system.

FOOTNOTE

1. Study is one of a series conducted by the Planned Change Project team and supported by the (national) Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison.

REFERENCES

1. Rogers, Carl R., "Toward a Theory of Creativity," *A Source Book for Creative Thinking*, S. J. Parnes and H. J. Harding, (eds.), Scribner's Sons, New York, 1962.
2. Rogers, Carl R., *On Becoming a Person: Therapists' View of Psychotherapy*, Houghton Mifflin, Boston, 1961.
3. Rogers, E. M., *Diffusion of Innovations*, Free Press of Glencoe, New York, 1962.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Conceptual Pre-Structuring for Detailed Verbal Passages[1]

**BARTON B. PROGER, RAYMOND G. TAYLOR, JR., LESTER MANN,
JOHN M. COULSON, ROBERT J. BAYUK, JR.**
Research and Information Services for Education, King of Prussia, Pennsylvania

ABSTRACT

In specific, meaningful printed passages various methods of conceptual organization can be used to increase students' understanding of key points of the passage. In a treatment by sex by ability level ANOVA, 124 Ss were randomly assigned to four types of "advance organizers": (a) Completion pretest (CP), (b) True–false pretest (TF), (c) Sentence outline (SO), and (d) Paragraph abstract (PA). On a 20-item multiple-choice posttest, designed to measure specific and general facilitative effects, three orthogonal comparisons for the two posttest part scores were made. Covert treatment groups (PA and SO) received significant specific facilitation; no differences among methods were found for general facilitation.

IN FORMAL learning situations, readers of highly specific verbal material frequently encounter difficulty in distinguishing between major and minor ideas. Authors of textbooks, among others, have sometimes tried to solve this problem by providing a summary overview of key concepts at the start of each new chapter, or unit of chapters, and by employing boldface type, italics, underlining, abstracts, and other organizing aids.

This study views three categories of conceptual organizing aids. The first category is "advance organization." Before a reading selection, for example, an abstract establishes a general conceptual background for the absorption of more specific details in the passage itself. Second, there is "concurrent organization." In this type, key words or even entire sentences embodying major concepts are underlined or italicized. The third category is "post organization." At the close of a chapter or unit of instruction in a textbook, for instance, key points are summarized.

A few isolated studies have been done in each of these areas of conceptual organization. In each of the three fields, little has been established regarding the actual mechanisms at work. For example, in the area of advance organization (conceptual pre-structuring), only the paragraph abstract method has been studied in any depth. Thus, at a time when much research in realistic settings is being reported, a pressing need for explicit guidelines exists because of the accelerating use of self-contained, study materials. If unfamiliar supplementary material is presented by self-contained, independent study packages, then the most efficient method of pre-organizing students' thoughts should be used. Results from controlled investigations of organizing mechanisms can be of immediate and widespread educational utility and can contribute also to knowledge of how verbal concepts are learned in realistic situations. Support for this position can be found in Ausubel (1), Ausubel and Fitzgerald (3, 4), Bloom and Goldstein (8), and Scandura and Wells (31). The present experiment studies not only the paragraph abstract type of advance organizer but also three other types with respect to both specific and general conceptual facilitation.

Ausubel (1) first postulated the basic rationale for phenomena involving conceptual organization. He hypothesized that both learning and retention could be aided by providing the reader with the key points of an unfamiliar but meaningful passage prior to its presentation ("advance organization"). He tried to bring about this conceptual structuring by using two different types of 500-word introductory passages. Ausubel reasoned major concepts presented in such introductory passages would be of such a level of generality and inclusiveness so as to promote the subsumption of more specific details presented later in the regular 2,500-word passage. With different levels of inclusiveness of concepts, he hypothesized there would result a hierarchical arrangement of concept subsumption.

To test his hypotheses, Ausubel (1) gave 120 college Ss a 500-word introductory passage twice for 5 minutes each two days before, and also immediately before, the complete selection. The same 2,500-word passage on the metallurgical properties of plain carbon steel was given to all Ss. However, prior to its presentation, two types of introductory passages ("advance organizers") were compared. The one type (*E*) embodied directly relevant key concepts about the steel-making process; the second type (*C*) contained only historical, non-process details about the steel-making procedure. The posttest consisting of twenty-six multiple-choice items was given immediately after the 20 minutes of main-passage reading. He found that $E > C$ $(.05 > P_{\text{1-tail}} > .01)$. In his discussion of results, Ausubel (1) claimed that the historical introduction did not provide any basis for the subsumption of more specific process-related details in the main passage itself. However, with the much more directly related introductory passage, Ausubel (1: 270) asserted, "The introduced subsumers thus became advanced organizers or anchoring foci for the reception of new material." It should be noted that Ausubel dealt only with the format of paragraph abstract for his advance organization treatments.

Next, Ausubel and Fitzgerald (3) studied advance organization in the light of conceptual discriminability theory. The reading passages used by Ausubel (1) were almost totally unfamiliar to the Ss. However, the investigators reasoned that in most learning situations Ss usually have some prior related knowledge. They hypothesized that the twofold role of advance organizers would be (a) to provide some conceptual background for new material and (b) to provide the reader with a basis for discriminating new concepts that might be confused with previously learned concepts. Ausubel and Fitzgerald decided therefore to investigate the effectiveness of different types of advance organizers in relation to material somewhat familiar to the Ss.

In attempting to choose material that had relevance to previously learned matter, Ausubel and Fitzgerald (3) used a passage on Buddhism (relying on its relationship to already acquired concepts about Christianity). Three types of advance organizers were chosen for use with the common Buddhism reading selection. The comparative organizer pointed out major differences and similarities between Buddhism and Christianity. The expository organizer was similar to the *E* treatment used in Ausubel (1): key facts about Buddhism were presented but were not contrasted with corresponding elements of Christianity. The historical organizer was similar to the control treatment in Ausubel (1); only historical background material was presented.

The Ss in Ausubel and Fitzgerald (3) were 155 college seniors. Each organizer of about 500 words was given 2 days before the 2,500-word Buddhism passage. The organizer was read for 8 minutes, while the Buddhism passage was read for 35 minutes. Alternate forms of a 45-item, multiple-choice posttest were given 3 and 10 days after the Buddhism passage. Significant differences on the Buddhism retention test in favor of the Ss with greater knowledge of Christianity were found. This result supported the reasoning of Ausubel and Fitzgerald that conceptual background for new concepts in terms of related but specifically different prior concepts, do facilitate learning by means of conceptual discriminability. As revealed by the posttest, only the comparative organizer was significantly different from the other two methods after 3 days, while on the posttest both the comparative and expository organizers were significantly different from the third treatment after 10 days. These results, however, pertain only to those Ss who were below the median on the Christianity test. In other words, advance organizers based on conceptual discriminability seem to work most effectively only with Ss whose conceptual background in the area of concern is relatively weak.

In the third relevant study, Ausubel and Fitzgerald (4) studied advance organization in the light of sequential learning. They were particularly interested in a sequential series of reading passages related to each other but not to previous student learning. In line with their previous reasoning, Ausubel and Fitzgerald (4) claimed advance organizers can provide conceptual background only, not increased conceptual discriminability. The Ss were 143 college seniors. *E* received a 500-word advance organizer directly related to the two reading selections to follow, while *C* received a passage which, while related, did not have nearly as close a relationship. Both *E* and *C* were allowed to study their respective advance organizers twice: for 8 minutes given 2 days before the first passage and for 6 minutes given immediately before the first passage. Both *E* and *C* received the same first passage (1,400 words with 25 minutes allowed) and the same second passage (1,600 words with 27 minutes allowed). Both passages dealt with the endocrinology of pubescence. A 36-item multiple-choice posttest was made for each passage. The test on the first passage was given 2 days after the reading of that passage, and the second passage was read 3 days later. The posttest on the second passage was given 4 days later.

Ausubel and Fitzgerald (4) found on the first posttest that $E > C$ $(P = .07)$. However, just as in Ausubel and Fitzgerald (3), they found that nearly all of this significant difference was caused by Ss in the bottom third in verbal ability. The investigators reasoned that advance organizers were effective mainly with poor verbal ability Ss because they initially possess poor conceptual organization, while high verbal ability Ss have the power to organize new concepts spontaneously. Finally, they found the advance organizer designed expressly for the first reading selection had no facilitative effect on the learning and retention of the second reading passage.

A fourth study, by Bloom and Goldstein (8), examined the results of five experiments in technical writing, one of which has relevance to the present study. The same Ss were used to compare two different presentation formats, each dealing with aircraft detection systems. *E* received an abstract of a longer, unabridged passage only; *C* received only the original, long passage. They claimed (8: 279):

> The abstracted version, though shorter, took comparatively longer to read (4 minutes for 572 words in the abstracted version, compared to 14.5 minutes for 3,100 words in the original). The key to this apparent contradiction lies in high 'information density'—an academic term defining content per given number of words. We found that the rapid reader suffers most from high information density.

A fifth study is that of Scandura and Wells (31), one of the most recent reported experiments on advance organization. The Ss were 104 undergraduate college students randomly assigned to one of four treatments. These four treatments were combinations of two types of mathematical material in the reading passages themselves (group theory and topological theory) and two types of advance organizers (historical introduction and mathematical-game, passage-related introduction). Each type of passage material had its respective 9-item criterion posttest. In each type of material, both the organizer and the passage itself were 1,000 words long. It was found that the combined mathematical-game organizer groups achieved better than the combined historical-introduction organizer groups ($P > .05$), and that this effect was stronger in the topology group than in the group theory group.

The role of the older organizer studies in a subsumption theory framework has been reviewed in depth by Ausubel (2). Popham (27) reviews some of the works briefly.

The purpose of the present experimental study was to identify and describe certain mechanisms of structuring that occur with highly specific, meaningful printed passages through the use of advance organizers. Previous relevant studies mentioned above have only considered variations on the paragraph abstract type of "advance organizer." In this study, four different types of advance organizers were compared: (a) paragraph abstract, (b) an enumerated sentence outline, (c) a true-false pretest, and (d) a completion pretest. Each of these pre-passage structuring methods dealt with the same set of ten key points from the reading passage employed as the learning task in the study.

Procedure

A highly detailed passage[3] on the nonreligious aspects of Amish customs and life was adapted from Hostetler (19). The passage consisted of about 2,600 words. The material, although filled with specific details, was easy to read and could be managed with a reading rate of about 200 words per minute. Maximum reading time of 15 minutes was thus allotted to the passage.

From this passage, ten broad, subsuming concepts were identified as the basis for each of four advance organization aids. Five twelfth-grade social studies classes from a suburban school district in Montgomery County, Pennsylvania, provided Ss. Each of the five classes were homogeneously grouped. The factor of classes, therefore, corresponded to ability levels. The five classes represented the whole range of ability in the twelfth grade in this school. Within each intact class, the four treatments were randomly assigned on the basis of a seating chart supplied by the teacher and administrator before the study began. The actual class sizes on the day of the experiment ranged from seventeen to twenty-nine. A total of 124 students were involved for the five classes.

On the day of the experiment, an investigator went to each of the five classrooms and remained there for the 50-minute duration. The first 5 minutes of the period were used to explain the purpose of the study. The Ss were assured that all results would remain anonymous. Instructions and one of the four advance organizer aids were appended to the front of the passage to be read.

After receiving general instructions, Ss of each group were told to process the advance organizer material carefully for 5 minutes according to the exact instructions printed on that sheet of the booklet. They were then told to read the passage itself carefully up to a 15-minute limit. After time was called, the advance organization sheet and the passage booklet were collected and at the same time the criterion test and IBM 1230 answer sheet were distributed.

The criterion test was of a 20-item multiple-choice format; each item had four options. The position of the correct option was randomly assigned within each test item. The items were arranged on the test according to the order in which the corresponding content matter originally appeared in the passage itself. The students were allowed 15 minutes for the test. The test items covered eight concepts which had been stressed in each of the advance organization methods (general, subsuming concepts) and twelve not covered by them (specific, subsumed concepts) but discussed in the reading passage itself. Subanalyses of the two parts of the test enabled generalized practice effects of advance organization to be separated from specific facilitative effects. Any differences found on the eight posttest items previously covered on the advance organizer can be attributed to specific conceptual facilitation; similarly, any differences found on the twelve concepts not stressed on the advance organizer can be attributed to general conceptual facilitation. Such subanalyses have seldom been done before on the effects of conceptual structuring mechanisms of the type used in this study. These subanalyses were necessary to the study's main concern: to compare the effectiveness of the combined "covert response" groups (paragraph abstract and sentence outline groups) with the combined "overt response" groups (true-false pretest and completion pretest groups).

It was hypothesized that student achievement in the former two groups (of covert response type) would be less than student achievement in the latter two groups (the overt response type). This expectation was based partially on the results of programmed instruction studies pitting overt response against covert response, e.g. Hershberger (18); Krumboltz and Weisman (22); Goldbeck and Campbell (15); Williams (37); Stolurow and Walker (34). The greater effectiveness of overt responses was also postulated on the basis of "arousal" theory in the sense of Hebb (17), Malmo (23), and Easterbrook (10). The two covert groups (paragraph abstract and sentence outline) had the advantage of "being given the answers" in a direct fashion. It was necessary to separate the criterion items into those specifically stressed on the advance organizers and those not so emphasized. Thus, the overt vs. covert response hypothesis demanded two separate analyses (specific facilitation effects and general facilitation effects).

Results

Two 4 × 2 × 5 (treatments by sex by ability level) analyses were run: one on the eight pretested items and one on the twelve non-pretested items. Since unequal cell frequencies were present, the usual analyses of variance had to be modified. Because the method of unweighted means (Winer, 38; Snedecor and Cochran, 33) has logical and computational difficulties, cell means were used as the units of analysis as though single replication had been used (Finney, 12; Edwards, 11). The expected mean squares were derived under the assumption of a completely fixed design. The highest order interaction thus became the estimate of experimental error.

For the eight pretested items (See Table 1), covert treatment groups (*PA* and *SO*) did significantly better ($P < .01$).

Boys achieved significantly higher than girls on these items ($P < .01$). The control factor of ability levels also resulted in significant differences ($P < .01$). Because of the significant treatment by ability level interaction ($.01 < P < .05$), an analysis for the simple effects of methods had to be carried out (See Table 2).

Table 1.—Summary of the 4×5×2 Analysis of Variance of Posttest Scores (8 Pretested Items)

Source	df	MS	F
Treatment (A)	3	3.04	20.83†
Sex (B)	1	1.04	7.10†
Ability (C)	4	2.02	13.82
A × B	3	.12	<1
A × C	12	.41	2.83*
B × C	4	.30	2.07
A × B × C	12	.15	

†p<.01.
*p<.05.

For the pretested items, the covert treatment groups did significantly better for lowest ability ($P < .05$), for low ability ($P < .01$), and for average ability ($P < .01$); however, no significant treatment differences could be found for the above average ability level and the highest ability level. There were no significant differences between the two overt response groups or between the two covert response groups in the eight "advance-organized" items. Figure 1 shows the nature of the significant treatment by ability level interaction.

General facilitative effects were studied by means of the non-"advance organized" criterion test items. No significant differences were found either for treatment or for sex; the control factor of ability level was significant at the .01 level. No interactions were significant.

Discussion

Significant results were found only for the eight pretested items (specific facilitation); the combined covert groups did significantly better than the combined overt groups. The results may

Table 2.— Summary of Analysis of the Simple Effects of Treatments for all Levels of Ability (8 Pretested Items)

Source	df	MS	F
A for C_1	3	.43	2.97
C_2	3	.14	<1
C_3	3	1.00	6.83†
C_4	3	2.26	15.50†
C_5	3	.81	5.57†
Error	12	.15	

†p<.01.
*p<.05.

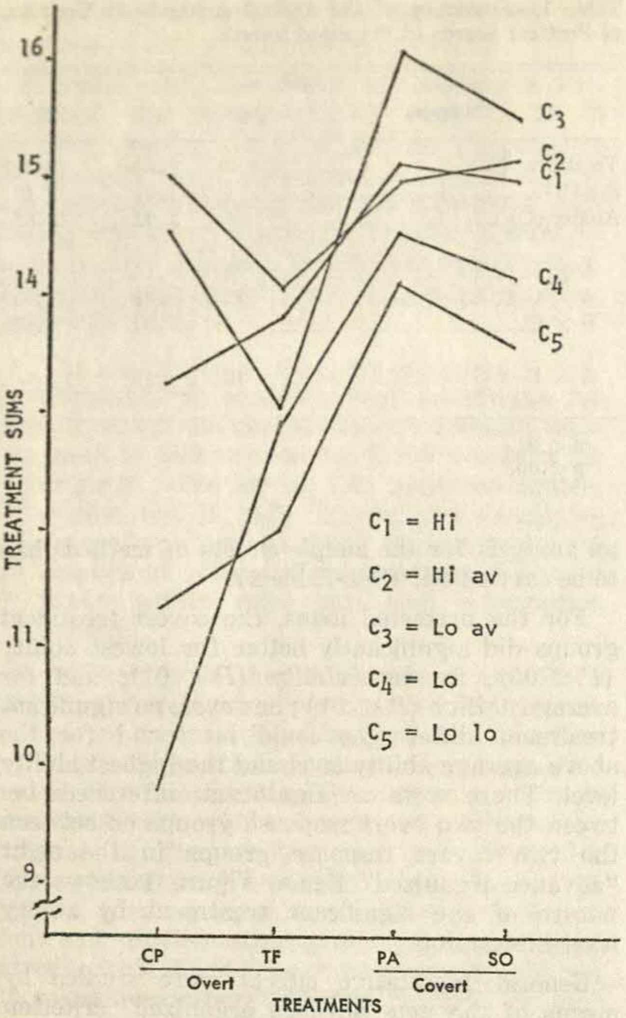


Figure 1. Simple Main Effects of Treatments for Ability Levels (8 Pretested Items)

be explained by the covert response groups (paragraph abstract and sentence outline) being more aware of the general organization of the passage in the sense of general drive activation (e.g., see Malmo, 23). Another explanation for the superior performance of the two covert response groups might be the "giving of answers" in the paragraph abstracts and sentence outlines. In any case, paragraph abstracts and sentence outlines are favored advance organizers for specifically stressed concepts. The interaction between treatments and ability levels acted in the same way as that found in Ausubel and Fitzgerald (3, 4). Evidently, specific conceptual facilitation occurs only for those Ss whose conceptual organization is initially poor.

None of the advance organizers showed any evidence for general conceptual facilitation; i. e., on the twelve test items not stressed on the advance organizers, no significant differences were found in treatments. Perhaps the singularly verbal nature of the paragraph abstract is not sufficient in and of itself for general facilitation. Triezenberg's (35) recent study has provided some evidence in this direction. He was interested in different degrees of abstraction in advance organizers: (a) verbal, (b) sketch, and (c) mechanical model. The verbal treatment involved only a verbal presentation; the sketch treatment involved the verbal presentation plus the sketch presentation; the mechanical model treatment involved the verbal presentation plus the mechanical model presentation. The subject matter was equilibrium in ecological systems. Besides three levels of methods, the factors of ability (low, medium, and high) and grade level (seventh and ninth) were studied. The criterion posttest used to measure immediate achievement and delayed achievement (6 weeks later) consisted of sixty multiple-choice items, equally divided among three areas (knowledge, comprehension, and application). For each combination of treatment group, ability level, and grade level, fifteen Ss were present in the final analysis sample. For the comprehension subtest, treatments were significant ($P < .05$); in particular, the model treatment was significantly different from each of the verbal and sketch treatments ($P < .05$ in each case). No significant treatment differences were found on the knowledge and application subtests. Thus, combining two media in an advance organizer treatment can be more effective than a single medium.

Recent advances in conceptual structuring theory have been tied to mathemagenic, or attention-producing, behavior in the sense of Rothkopf (30). Advance organization studies in this new context are those of Gustafson and Toole (16) and Carter and others (9). A revival of conceptual discriminability as a variable in advance organization has been reported by Ausubel, Stager, and Gaite (5), Gaite, Ausubel, and Stager (14), and Shuell and Hapkiewicz (32). Concurrent organization of a mathemagenic type has been studied by Frase (13), Pyper (29), and Natkin and Stahler (25). Post organization results were reported by Bauman and Glass (6). Additional recent work on the use of pictures as organizers for learning from printed discourse has been completed by Koenke (20) and Koenke and Otto (21). Finally, advance organization has even been studied with the mentally retarded: Blackhurst (7) and Neisworth (26).

Further research is needed in the field of different types of advance organization. Other than the present study, most experiments have investigated only the paragraph abstract type of advance organizer and have not gone into the general and specific facilitative effects. Another virtually untapped area for research is the possibility of interaction between advance organization and concurrent organization. For example, one might compare the four types of advance organizers of the present study in the presence and absence of a particular type of concurrent organization (perhaps underlining of key words and concepts in the passage itself). Proger and others

(28) have completed such a study. The implications these studies have for independent study materials are manifold. Nonetheless, far from being a specialized topic within educational psychology, the need for improved conceptual structuring mechanisms is being felt in noneducational circles (Mathes, 24; Weil, 36). Clearly, the psychology of subsumption theory is growing rapidly.

FOOTNOTES

1. This study was supported by Research and Information Services for Education, an ESEA Title III agency (OEG #1-7-673010-2696). It was also partially supported by the School District of Springfield Township, Oreland, Pennsylvania. However, no endorsement of the practices contained herein on the part of either or both supporting agencies should be inferred. This study represents a greatly expanded and revised version of a paper presented at the 77th annual convention of the American Psychological Association in Washington, D. C., August 31, 1969.
2. The investigators wish to acknowledge the generous support and cooperation given to the study by the personnel of Springfield School District, especially to Dr. Steven A. Kalapos, Assistant to the Superintendent for Instruction and to the other personnel involved in the study: F. G. Judd, J. J. Gillespie, A. C. Eickhoff, J. Ragan, and W. Will—all of Springfield School District. Finally, we wish to express our appreciation to H. Wolfenberger of the Graduate Center Library, Pennsylvania State University, King of Prussia Branch, and to D. N. Spaans of Research and Information Services for Education.
3. Permission to use and adapt the copyrighted Amish material has been obtained from the Mennonite Publishing House, 610 Walnut Avenue, Scottdale, Pennsylvania.

REFERENCES

1. **Ausubel, D. P.**, "The Use of Advance Organizers in the Learning and Retention of Meaningful Material," *Journal of Educational Psychology,* 51: 267–272, 1960.
2. **Ausubel, D. P.**, *Educational Psychology: A Cognitive View,* Holt, Rinehart, and Winston, New York, 1968.
3. **Ausubel, D. P.; Fitzgerald, D.**, "The Role of Discriminability in Meaningful Verbal Learning and Retention," *Journal of Educational Psychology,* 52: 266–274, 1961.
4. **Ausubel, D. P.; Fitzgerald, D.**, "Organizer, General Background, and Antecedent Learning Variables in Sequential Verbal Learning," *Journal of Educational Psychology,* 53: 243–249, 1962.
5. **Ausubel, D. P.; Stager, M.; Gaite, A. J. H.**, "Proactive Effects in Meaningful Verbal Learning and Retention," paper presented at the meeting of the American Educational Research Association, Los Angeles, February, 1969.
6. **Bauman, D. J.; Glass, G. V.**, "The Effects on Learning of the Position of an Organizer," paper presented at the meeting of the American Educational Research Association, Los Angeles, February, 1969.
7. **Blackhurst, A. E.**, "Auditory Learning of Retarded Adolescents in a Subsumption Theory Context," unpublished doctoral dissertation, University of Pittsburgh, 1966.
8. **Bloom, J. H.; Goldstein, S. A.**. "Five Experiments in Technical Writing," in Estrin, H. A. (Ed.) *Technical and Professional Writing: a Practical Anthology,* Harcourt, Brace and World, New York, 1963, pp. 278–281.
9. **Carter, C. E., Jr.; Proger, B. B.; Mann, L.; Haughey, C. F.; Morris, V. R.; Reckless, D. E.**, "Advance Conceptual Structuring Mechanisms of Overt and Covert Response Types in Detailed Reading Passages," paper presented at meeting of American Educational Research Association, Minneapolis, March, 1970.
10. **Easterbrook, J. A.**, "The Effect of Emotion on Cue Utilization and the Organization of Behavior," *Psychological Review,* 56: 183–201, 1959.
11. **Edwards, A. L.**, *Experimental Design in Psychological Research* (Revised edition), Holt, Rinehart, and Winston, New York, 1960.
12. **Finney, D. J.**, *An Introduction to the Theory of Experimental Design,* University of Chicago Press, Chicago, 1960.
13. **Frase, L. T.**, "Effective Prose Reading: Shaping and Discriminative Effects of Questions," paper presented at meeting of American Educational Research Association, Chicago, February, 1968.
14. **Gaite, A. J. H.; Ausubel, D. P.; Stager, M.**, "Retroactive Facilitation in Meaningful Verbal Learning: Generality over Time and Effect of Differing Levels of Verbal Ability," paper presented at meeting of American Educational Research Association, Los Angeles, February, 1969.
15. **Goldbeck, R. A.; Campbell, V. N.**, "The Effects of Response Mode and Response Difficulty on Programmed Learning," *Journal of Educational Psychology,* 53: 110–118, 1962.
16. **Gustafson, H. W.; Toole D. L.**, "Effects of Adjunct Questions, Pretesting, and Degree of Student Supervision on Learning from an Instructional Text," paper presented at meeting of American Educational Research Association, Los Angeles, February, 1969.
17. **Hebb, D. O.**, "Drives and the C. N. S. (Conceptual Nervous System)," *Psychological Review,* 62: 243–254, 1955.
18. **Hershberger, W. A.**, "Self-evaluation Responding and Typographical Cueing: Techniques for Programming Self-instructional Reading Materials," *Journal of Educational Psychology,* 55: 288–296, 1964.
19. **Hostetler, J. A.**, *Amish Life,* Herold Press, Scottdale, Pennsylvania, 1959.
20. **Koenke, K.**, "The Roles of Pictures and Readability in Comprehension of the Main Idea of a Paragraph," paper presented at meeting of American Educational Research Association, Chicago, February, 1968.
21. **Koenke, K.; Otto, W.**, "Contribution of Pictures to Children's Comprehension of the Main Idea in Reading," *Psychology in the Schools,* 6: 298–302, 1969.
22. **Krumboltz, J. D.; Weisman, R. G.**, "The Effect of Intermittent Confirmation in Programmed Instruction," *Journal of Educational Psychology,* 53: 250–253, 1962.
23. **Malmo, R. B.**, "Activation: a Neurophysiological Dimension," *Psychological Review,* 66: 367–386, 1959.
24. **Mathes, R. C.**, " 'D' People and 'S' People," *Science,* 164: 630, 1969.
25. **Natkin, G.; Stahler, E.**; "The Effects of Adjunct Questions on Short and Long-Term Recall of Prose Materials," *American Educational Research Journal,* 6: 425–432, 1969.
26. **Neisworth, J. T.**, "Use of Advance Organizers with the Mentally Retarded," unpublished doctoral dissertation, University of Pittsburgh, cited in Durkee, F. M., *Thirteenth Annual Report of Educational Research: Pennsylvania Colleges and Universities,* Pennsylvania Department of Education, Harrisburg, Pennsylvania, 1969, p. 166.
27. **Popham, W. J.**, "Curriculum Materials," *Review of Educational Research,* 39: 319–338, 1969.

28. **Proger, B. B.; Mann, L.; Taylor, R. G., Jr.; Morris, V. R.; Reckless, D. E.,** "Advance and Concurrent Conceptual Structuring Mechanisms in Detailed Reading Passages,"paper presented at meeting of American Educational Research Association, Minneapolis, March, 1970.
29. **Pyper, J. R.,** "Motivation and Mathemagenic Behaviors: Toward a Theory of Inserting Questions in Instructional Material," paper presented at the meeting of the American Eucational Research Association, Los Angeles, February, 1969.
30. **Rothkopf, E. Z.,** "Learning from Written Instructive Material: An Exploration of the Control of Inspection Behavior by Test-like Events," *American Educational Research Journal,* 3: 241–249, 1966.
31. **Scandura, J. M.; Wells, J. N.,** "Advance Organizers in Learning Abstract Mathematics," *American Educational Research Journal,* 4: 295–301, 1967.
32. **Shuell, T. J.; Hapkiewicz, W. G.;** "Similarity and Retroaction in Prose Material," paper presented at meeting of American Educational Research Association, Los Angeles, 1969.
33. **Snedecor, G. W.; Cochran, W. G.,** *Statistical Methods* (6th edition), Iowa State University Press, Ames, 1967.
34. **Stolurow, L. M.; Walker, C. C.,** "A Comparison of Overt and Covert Response in Programmed Learning," *Journal of Educational Research,* 55: 421–429, 1962.
35. **Triezenberg, H. J.,** *The Use of an Advance Organizer in Teaching Selected Concepts of Ecological Systems,* Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, February, 1968, Technical Report No. 42, Center No. C-03, Contract OE 5-10-154.
36. **Weil, B. H.,** "Abstracts with Zest," *Science,* 165: 6, 1969.
37. **Williams, Joanna P.,** "Comparison of Several Response Modes in a Review Program," *Journal of Educational Psychology,* 54: 253–260, 1963.
38. **Winer, B. J.,** *Statistical Principles in Experimental Design,* McGraw-Hill, New York, 1962.

Research Briefs

A STUDY OF DOGMATISM AND ANXIETY IN ELEMENTARY SCHOOL CHILDREN

THIS study was undertaken for two purposes as follows: (1) to determine the relationship between dogmatism as measured by an elementary school form of Rokeach's Dogmatism Scale (ESFDS) and manifest anxiety as measured by the Children's Manifest Anxiety Scale (CMAS) and (2) to determine if there was a difference between high dogmatic children (HDC) and low dogmatic children (LDC) on measures of manifest anxiety. Both of the scales, i.e. ESFDS and the CMAS, were developed from prototype scales used primarily with adult subjects. The ESFDS scale was developed from Rokeach's Dogmatism Scale (2) and the CMAS scale was derived from Taylor's Anxiety Scale (1). Inasmuch as previous investigations had not been done with children; this experimenter thought such data might prove a useful supplement to existing data obtained from adult populations (3).

Subjects and Procedures

Sixth-grade students ($N = 160$) enrolled in ten elementary schools were selected at random from the total sixth-grade population of an all white, lower middle-class community. The students were administered the ESFDS and the CMAS scales. A Pearson product-moment correlation was computed to determine the relationship between dogmatism and manifest anxiety. On the basis of the ESFDS scores, the students were grouped into HDC (upper quartile) and LDC (lower quartile) groups. A single classification of the analysis of variance (ANOV) was computed to determine differences in these two groups.

Results

The correlation between ESFDS scores and CMAS scores was .36. A t-test indicated that this relationship was significant ($p < .001$). A t-test computed for the ANOV indicated that the HDC group was significantly ($p < .001$) more anxious than the LDC group.

Conclusion

While the relationship between dogmatism and manifest anxiety in children is small, it is a highly significant relationship. The correlational aspects of this study yielded results very similar to those found in Rokeach's correlational studies with adults (3).

Children who were closed-minded (HDC) exhibited significantly more manifest anxiety than the children who were open-minded (LDC). According to Rokeach's definitions for open-closed-mindedness (4), one can see why a child with a high potential for anxiety (HDC) would have difficulty making cognitive distinctions between data source and authority.—**Edward P. Merryman, Ball State University, Muncie, Indiana.**

References

1. **Castaneda, A., McCandless, B. R., and Palermo, D.,** "The Children's Form of the Manifest Anxiety Scale," *Child Development,* 27: 317–326, 1956.
2. **Figert, R.,** "An Elementary School Form of the Dogmatism Scale," unpublished doctor's dissertation, Teachers College, Ball State University, Muncie, Indiana, 1965.
3. **Rokeach, M.,** *The Open and Closed Mind,* Basic Books, Inc. Publishers, New York, 1960.
4. **Rokeach, M.,** "In Pursuit of the Creative Process," in *The Creative Organization,* Steiner, Gary A., (ed.), University of Chicago Press, Chicago, 1965.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 1, September 1970)

Use of the Canonical Correlation Analysis in an Investigation of Pupil's Rate of Learning in School[1]

MARGARET C. WANG, University of Pittsburgh

ABSTRACT

This study investigated the relationship between measures of pupil's rate of learning and some selected pupil aptitude and achievement measures as well as certain classroom performance measures. It was carried out in the context of the Individually Prescribed Instruction Project (IPI) of the Learning Research and Development Center, University of Pittsburgh. Ss were 629 elementary school students who worked in six different units of IPI mathematics between September 1967 and January 1968. The general hypothesis was tested through the use of Pearson product-moment correlation analysis, Multiple correlation analysis, as well as canonical correlation analysis. The overall results seem to substantiate that it is possible and appropriate to use the Canonical Correlation Analysis to test the research hypothesis formulated, and the selected variables used are related to how fast a particular student learns the material under an individualized instructional program. However, because of the fluctuation of the relative importance of the predictors from unit to unit as indicated by the results of Multiple Correlation Analysis as well as the canonical correlation analysis, the results did not offer a clear cut explanation about the nature of the relative contributions made by each of the predictors to the predictions of the rate measures. It may be concluded that rate of learning is specific to a given task and is not a general factor characterizing student performance in all learning situations.

THE BASIC purpose of the study was to identify a useful measure of pupils' learning in school and to investigate the relationship between measures of pupils' rate of learning and some selected pupil aptitude and achievement measures as well as certain classroom performance measures. Research findings of various research studies on school learning have indicated that (a) there are substantial differences in levels of achievement and in learning rates among pupils in school learning situations, as well as under experimental laboratory conditions (1, 10, 11), and (b) these individual differences are a function of the characteristics of the learner and many other variables that are closely related to the learning task and the learning environment (2, 5, 9).

It is generally agreed both by educational theorists and practitioners that there are a host of factors affecting the pupil's learning performance. Pupil's level of achievement and the speed with which he learns depend upon the unique individual characteristics that the pupil brings with him to the learning situation, the nature of the task to be learned, and the conditions under which the particular learning takes place. This recognition of the multiplicity of factors that influence the pupil's level of achievement, as well as differences in the amount of time each pupil needs to achieve mastery, have focused the emphasis on the importance of making provisions for individual differences among pupils in our schools.

Because of this changing trend in the instructional procedures and the emphasis on the individualization of instruction, the methods of assessing pupil's learning performance also need to be modified. Under the "conventional classroom situations," instructional procedures are planned for the classroom group as a whole; pupils are to learn the particular lessons in a fixed time allowance, and everyone is to proceed at essentially the same pace to reach the goals. In other words, possible individual differences in the rate of progress are not taken into consideration; a pupil's success or failure in school learning is judged in terms of the amount or the degree he has achieved when learning time is held constant.

However, under the instructional programs where individualization is provided, the level of achievement cannot be used as an assessment of pupil's learning performance, because the level of achievement is held constant for every pupil who learns the particular task. The individual differences are indicated by the speed at which each pupil learns the specific task in the learning sequence of the instructional objectives. Hence, under the individualized instructional programs, the rate of learning is considered as an achievement measure—an assessment of the pupil's learning performance.

This investigation was undertaken to study the characteristics of "rate of learning" by exploring its relationships with various factors that are hypothesized as being related to rate. Specifically, the study proposed to (a) identify a measure or measures in evaluating pupil's learning performance under individualized instructional programs, (b) specify the independent predictors that are associated with the rate measures, (c) determine the relative contributions made by each of these predictors to the prediction of the criterion measure of rate and (d) identify the interrelationships among the predictors—factors that are associated with learning rate.

To investigate problems in the measurement of rate of learning and identify factors that are associated with differences in rate, it is necessary to work in situations where variations in rate actually exist. For this reason, the study was carried out in the context of the Individually Prescribed Instruction Project (IPI) of the Learning Research and Development Center at the University of Pittsburgh.

The IPI procedure is a program designed to achieve a certain type of individualized instruction in Grades K through 6 in the subject areas of reading, arithmetic, and science. Its basic elements included (a) detailed sequences of behavioral objectives which define the abilities that each pupil is to acquire, (b) study materials, that are largely self-instructional in nature, to teach each objective, (c) a testing program for placing each pupil at the proper point in the curriculum sequence and for monitoring his progress, and (d) classroom management procedures that permit each pupil to proceed at a rate best suited to his needs. In mathematics, the curriculum is organized in terms of units (numeration, place value, addition, subtraction, multiplication, division, combination of processes, fractions, money, time, system of measurement, geometry, and special topics) and levels (A, B, C, D, E, F, G, H, and I). This means that a pupil will typically work through the series of topics at one level before moving on to the next. The work in a given topic at a given level, such as Level E-Numeration or Level C-Addition, is identified as a unit and is defined on the basis of a certain number of objectives. The objectives are identified as skills. Study materials have been developed to enable the pupil to master these objectives and are identified as work pages. (For detailed descriptions of IPI Program see (8).)

The research population of this study consisted of 629 elementary school students who worked in six different units of IPI mathematics between September 1967 and January 1968. The samples are drawn from two schools in which individually prescribed instruction is used.

Four different measures of rate were used as criterion variables. The decision to use four different measures of pupil's rate of learning was based on findings from an earlier study (12) which indicated that any one measure had certain obvious restrictions in terms of describing a pupil's rate of progress. The four measures are described in detail by Wang (13). In brief, they are defined as follows:

$$\text{Rate } 1 = \frac{100 - \text{pretest score,}}{\text{days worked on unit}}$$

$$\text{Rate } 2 = \frac{\text{number of pages worked,}}{\text{days worked on unit}}$$

$$\text{Rate } 3 = \frac{\text{number of skills learned,}}{\text{days worked on unit}}$$

Rate 4 = total number of skills learned during experimental period.

As variables that should be related to rate, this study has investigated certain measures of the pupil's scholastic aptitude, academic achievement, and certain classroom performance characteristics that past work had shown to be related to learning. The nine measures used are: (a) verbal IQ, (b) non-verbal IQ, (c) mental age, (d) mathematics achievement score, (e) reading achievement score, (f) total number of skills worked during the previous school year, (g) attention score, (h) question$_1$, student's reaction to the difficulty level of the work, (i) question$_2$, student's reaction to the type of learning material used.

In order to examine these relationships as they are found with different types of specific mathematic content, student performance in six different units of IPI mathematics was investigated. As the first step in the assessment of the functioning of the measures involved, Pearson product–moment correlations were computed between each individual predictor and each individual rate measure for the six selected E-level units of the IPI mathematics. Tables 1–6 summarize the results for each of the E-level units. Only significant correlation coefficients are listed.

In looking at the significant Pearson product–moment correlation coefficients, one can see that none of the predictors is correlated with one rate measure or another in one or several of the units, these predictors are not found to be related generally to all rate measures in all situations. In most cases, the correlations between the rate measures and the predictors for each unit are not significantly different from zero. Among 216 correlations, only twenty-eight were significant. Even those coefficients that were statistically significant were quite small in absolute value; furthermore, the size of these coefficients was not at all consistent from unit to unit. This indicated that any single predictor would be of little use in predicting any one of the rate measures.

The results clearly support our research hypothesis that no single rate measure nor any single

Table 1.—Product-Moment Correlation Coefficient Between the Rate Measures and the Predictors for the E-Numeration Unit of the IPI Mathematics Curriculum (N=182)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math Achiev. (X_4)	Read. Achieve. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Rate$_1$									
Rate$_2$						.193**			
Rate$_3$									
Rate$_4$							.165*	.183**	.161*

(Only significant correlations are listed)

**.01 level of significance, critical R = .181, d.f. = 200
*.05 level of significance, critical R = .138, d.f. = 200

predictor can explain the complex nature of rate of school learning. For example, in Table 1 none of the predictors correlated significantly with rate$_1$, only one variable (the total number of skills completed during the previous school year) is correlated with rate$_2$ at .01 level of significance. Out of the thirty-six correlations for the numeration unit, only two were significant at .01 level and two at the .05 level.

One of the implications of Yeager's study (11) was that his finding of a consistent lack of correlation between predictors and rate measures was due to a lack of comprehensiveness in any single measure. This study investigated the possibility that individual prediction might be more highly related to rate if a more comprehensive rate index were used, the multiple correlation of each predictor with a composite of the four rate

Table 2.—Product-Moment Correlation Coefficient Between Each Rate Measure and Each Predictor for the E-Place Value Unit of the IPI Mathematics Curriculum (N=109)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math. Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques. (X_9)
Rate$_1$			.311**	.345**	.234*	.245*			
Rate$_2$						.335**			
Rate$_3$									
Rate$_4$									

(Only significant correlations are listed)

**.01 level of significance, critical R = .254, d.f. = 100
*.05 level of significance, critical R = .195, d.f. = 100

Table 3.—Product-Moment Correlation Coefficient Between the Rate Measures and the Predictors for E-Addition Unit of the IPI Mathematics Curriculum (N=42)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math. Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Rate$_1$					.386*	.403**	.412**		
Rate$_2$						.314*			
Rate$_3$									
Rate$_4$									

(Only significant correlations are listed)

**.01 level of significance, critical R = .393, d.f. = 40
*.05 level of significance, critical R = .304, d.f. = 40

measures was determined. The results of this multiple correlation analysis show that, as would be expected, the correlation between each predictor variable and the rate measures increase when all four rate measures are combined. For example, in the E-Numeration unit the multiple R for question$_1$ and the four rate measures combined is .247 which is larger than .183, which is the only significant correlation between question$_1$ and the single rate measure found in Pearson Product–Moment Correlation Analysis.

Among the multiple R's tested for each predictor and the four rate measures, the majority of them were not significantly different from zero. For example, for the numeration unit, only question$_1$ is significantly related to the rate measures. This general lack of an increase in number of significant correlations with single predictors when a composite rate measure is used suggests that the lack of relationship between predictors and rate is not due primarily to the lack of comprehensiveness in the rate measure.

Table 4.—Product-Moment Correlation Coefficient Between the Rate Measures and the Predictors for E-Subtraction Unit of the IPI Mathematics Curriculum (N=103)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math. Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Rate$_1$				.223*	.219*	.255**			
Rate$_2$		.242*				.341**			
Rate$_3$			.217*		.317**	.247**		.233*	
Rate$_4$							.260**		

(Only significant correlations are listed)

**.01 level of significance, critical R = .254, d.f. = 100
*.05 level of significance, critical R = .195, d.f. = 100

Table 5.—Product-Moment Correlation Coefficient Between the Rate Measures and the Predictors for the Multiplication Unit of the IPI Mathematics Curriculum (N=111)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math. Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Rate$_1$	.215*								
Rate$_2$									
Rate$_3$						.293**			
Rate$_4$									

(Only significant correlations are listed)

**.01 level of significance, critical R = .254, d.f. = 100
*.05 level of significance, critical R = .195, d.f. = 100

After analyzing the relationship between measures and each of the predictors, the correlation between each rate measure and all of the predictors combined was investigated. These multiple R's are shown in Table 7. A comparison of the correlation coefficients obtained from the simple correlation analysis and the multiple correlation analysis showed that all the multiple R's are larger than the Pearson product-moment r's, and a much larger percent of the multiple R's are found to be significant.

This result supports the hypothesis that the predictability of rate measures increase when all the predictors are included in the regression analysis. For example, the Pearson product-moment correlation coefficient between rate$_4$ and the predictors (verbal IQ, nonverbal IQ, MA, Mathematics achievement, reading achievement, skills, attention, question$_1$ and question$_2$) for the numeration unit are: .054, .069, .124, .99, .068, .023, .165, .183, and .161 respectively; the multiple R for rate$_4$ and all the predictors combined for

Table 6.—Product-Moment Correlation Coefficient Between the Rate Measures and the Predictors for the E-Combination of Process Unit of the IPI Mathematics Curriculum (N=69)

	Verb. I.Q. (X_1)	Non-Verb. I.Q. (X_2)	M.A. (X_3)	Math. Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Rate$_1$				.284*					
Rate$_2$					.262*				
Rate$_3$									
Rate$_4$						.288*			

(Only significant correlations are listed)

**.01 level of significance, critical R = .302, d.f. = 70
*.05 level of significance, critical R = .233, d.f. = 70

Table 7.—Beta Coefficients for Multiple Regression Equations for Predicting Indicated Rate Measure in Mathematics Units for Cases Where the Multiple Correlation is Significant

Rate Measure	Unit	Multi R	Verb I.Q. (X_1)	Non-verb I.Q. (X_2)	M.A. (X_3)	Math Achiev. (X_4)	Read Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.[1] (X_8)	Ques.[2] (X_9)
1	Place Value	.433**	-.093	-.024	-.172	.247	.022	.189	-.059	-.049	-.033
1	Addition	.642*	.325	-.494	.131	-.122	-.203	.353	.292	.081	-.217
1	Subtraction	.524**	.219	-.467	.219	.113	.086	.171	.159	.194	-.056
2	Numeration	.343**	.213	-.133	-.077	-.027	-.194	.280	-.049	-.113	-.038
2	Place Value	.450**	.016	-.132	-.040	.193	.044	.371	-.032	-.209	-.004
2	Subtraction	.502**	-.025	-.380	.275	.008	.033	.243	.205	.048	.083
3	Subtraction	.517**	.187	-.370	.247	.107	.263	.138	.035	.290	-.084
3	Multiplication	.416*	.064	.037	.072	.210	-.046	.241	.156	-.099	.026
4	Numeration	.337**	-.024	-.139	.265	.073	-.029	-.041	.213	.166	.133

**.01 level of significance
*.05 level of significance

the same unit is .337. The Pearson r between rate$_2$ and the predictors are —.048, —.112, —.088, —.066, —.119, .193, —.025, —.131, and —.125 respectively, while the multiple R for rate$_2$, and all the predictors combined for the same unit is .343.

Nevertheless, among the multiple R's tested for each rate measure and all the predictors for the selected E-level units of the mathematics curriculum, many are not significantly different from zero. For example, for rate$_1$ and rate$_2$ only 50 percent of the units studied showed significant multiple R's; for rate$_3$, 33 percent of the units had significant multiple R's. Therefore, among the 24 multiple R's, nine are significant—37.5 percent.

In those cases where the multiple correlation coefficients are significant, it is of interest to examine the relative contribution of the various predictors. One way of interpreting the individual contribution of the predictor variables is by examining the beta weights, the partial regression coefficients. The contribution of most predictors in this study can be expected to be quite small, since the multiple R's are relatively small. (See Table 7.)

Because of the small size of the betas and because of their known inconsistency from sample to sample when the sample size is small, it is not surprising to note the lack of any consistent pattern in the betas from one regression equation to another. This problem can be investigated further by examining the structure R's for these same relationships. These structure R's indicate the correlations between the original predictors and the derived linear composite of the predictors. The structure R's are more stable in that they do not fluctuate as much from sample to sample as the betas do. For this reason, it is more meaningful to use the structure R's to examine the relative contribution of the various predictors for interpretative purposes. The structure R's for the significant multiple correlations are listed in Table 8.

The lack of consistency from one regression equation to another in the size of these structure R's, as well as in the betas, suggests that the relative importance of the various predictors in determining rate of learning probably varies greatly from unit to unit and from one rate measure to another.

The results of the simple correlation analysis and the multiple correlation analysis indicate that no one predictor variable is a major determiner of rate of pupil learning, that rate is determined by a complex of factors, and that the relative importance of each of factor varies greatly from one learning task to another. In order to obtain a more complete assessment of pupil's learning in school under the IPI program, more than one measure of rate of learning should be used. This general hypothesis was examined further in the canonical correlation analysis.

Canonical correlation analysis was used to answer the research question—*What is the relationship between the best linear combination of the predictors and a linear composite of the rate measures?* The canonical correlation method is one of the correlational techniques. Simple product-moment correlation method is used when one wants to study the relationship between two variables, the predictor and the criterion variable; multiple correlation method is used when one wants to investigate the relationship between a criterion variable and two or more predictor variables. When one wishes to learn the relationship between more than one variable in the criterion set and more than one variable in the predictor set, the appropriate procedure to use is the canonical correlation method. For two sets of variables can contain very similar information even though pairs of variables selected one from each set may not be particularly highly correlated. This is because of the fact that what one variable is measuring in one set may be measured by a combination of variables in the other set or vice versa. Therefore, the underlying traits may be reflected in several variables; yet only when these several variables are examined in combination are the relationships between two sets of variables made evident.

As Kendall (7) pointed out, through the canonical correlation method, one may reduce the relationship between the two sets of variables to the simplest form and hence, enable the researcher to look at the interdependence between two sets of variables. The canonical correlation provides us with canonical correlation coefficients, which show us the degree of relationship between such linear composites (the canonical variates), and also provides indices of the relative contribution of each of the original variables to the composite of which it is a part.

The number of canonical correlations is equal to the rank of the intercorrelation matrix among the two sets of variables. For example, the rank of the matrix for the present investigation is four, the number of rate measures. Cooley and Lohnes (3) pointed out that it is algebraically impossible to have more non-zero canonical correlations than the number of variables in the smaller of the two sets. In general the complete factor structure of a set of variables contains as many factors as there are variables. Hence, it is obvious that if the larger set is composed of nine variables as in the case of this study, and the smaller set is composed of four variables, only four factors can be extracted from the smaller set. As a result, canonical R's are available between four of the factors in the larger set and the four factors of the smaller set. The remaining factors in the predictor set have no counterpart in the criterion set and do not enter into the canonical solution.

According to Hotelling (6) a canonical solution can be explained in terms of factor analysis—the two sets of variables are factor analyzed independently first and then develop weights which would rotate the two factor structures to maximum correlation. In the case of canonical analysis, the factors extracted (the linear combination of the predictor set and the linear combination of the criterion set) are called canonical variates, and

Table 8.—Structure R's of Each Variable for the Indicated Rate Measure in Mathematics Units for Cases Where the Multiple Correlation is Significant

		Structure R								
Rate Measure	Unit	Verb. I.Q. (X_1)	Non-verb I.Q. (X_2)	M.A. (X_3)	Math Achiev. (X_4)	Read. Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
1	Place Value	.084	.150	.718	.796	.539	.567	-.272	.141	-.112
1	Addition	.230	-.077	.304	.208	.602	.628	.643	.075	.110
1	Subtraction	.146	-.265	.292	.426	.418	.487	.384	.384	-.024
2	Numeration	-.139	-.328	-.257	-.193	-.349	.562	-.073	-.383	-.363
2	Place Value	-.105	-.261	.221	.388	.257	.743	-.131	-.286	-.334
2	Subtraction	-.304	-.483	.125	.185	.177	.679	.366	.088	-.302
3	Subtraction	.246	-.091	.420	.230	.613	.477	-.014	-.450	-.013
3	Multiplication	.594	.081	.369	.038	-.157	.484	-.166	.132	.181
4	Numeration	.160	.204	.369	.293	.203	-.068	.489	.543	.479

Table 9.—The Canonical Vectors (Beta Weights) Associated with the First Canonical Variates of Each of the Selected E-Level Units of the IPI Mathematics Curriculum

	Numeration	Place Value	Addition	Subtraction	Multiplication	Comb. of Processes
Criterion Set						
Rate$_1$	.124	.619	-.780	.577	-.062	-.041
Rate$_2$	-.752	.476	-.451	.461	-.471	.111
Rate$_3$	-.072	.191	-.062	.339	-.730	.658
Rate$_4$	.710	.071	.177	.162	-.233	.651
Canonical R	**.4383	*.4855	.6897	**.6562	*.4738	.5292
Predictor Set						
Verb. IQ	-.446	-.128	-.741	.234	-.233	-.727
Non-verb IQ	.025	-.160	1.081	-.876	-.033	.169
MA	.596	.286	-.497	.528	-.049	.228
Math Achieve	.198	.519	.133	.053	-.619	-.155
Read Achieve	.251	.067	.136	.266	.258	.518
Skills 1967	-.528	.622	-.590	.343	-.624	.696
Attention	.433	-.057	-.424	.258	-.272	-.088
Question$_1$	.475	-.258	.153	.399	.247	.404
Question$_2$	.364	-.003	-.073	-.137	-.164	.409

**.01 level of significance
*.05 level of significance

the correlation between the first canonical variate (factor) of the criterion set and first canonical variate of the predictor set is the first canonical correlation. Therefore, for the present study, by using the canonical correlation analysis technique, two new scores were developed through the combination of the four rate measures and the combination of the measures of student characteristics. The derived combination (factors), one from each set of variables, are the canonical variates. The correlation between the first two canonical variates (one from each set) is the first canonical correlation. This canonical correlation indicates the similarities between the two sets of variables.

The first canonical variate for the rate measures was obtained from weighing the pupil's original rate scores by the corresponding coefficients, and the first canonical variate of the predictor variables was obtained from weighing the pupil's original student characteristics scores by the corresponding coefficients. Each of these two derived first canonical variates are uncorrelated with the other derived canonical variates (other combination of the variables) in it's own set, but each has maximum correlation with its paired variate from the other set (4).

The canonical correlations were tested for significance by a chi-square approximation. In general, the results of the canonical analysis for each of the six units selected for this study indicate that there is a significant relationship between the domain of rate measures and the domain of the predictors.

Among the first canonical correlations of the six units, four are significant. Although the first canonical R's of the Additions and the Combination of Process unit were not significant, the size of the canonical R's was quite substantial, .656, and .549 respectively. The small sample size used in the analysis of these two units may have been the reason for the conflicting results.

The nature of these pairs of significant canonical variates can be inferred by examining the vectors (the partial regression coefficients) associated with them. (See Table 9.) For example, the first canonical variates for the numeration unit can be expressed as equations as follows:

First canonical variate for the criterion set =

$$.124z_{rate_1} - .752z_{rate_2} - .072_{rate_3} + .710z_{rate_4}$$

First canonical variate for the predictor set =

$$-.446z_{x_1} + .025z_{x_2} + .025z_{x_2} = .596z_{x_3} + .198z_{x_4} + .251z_{x_5} - .528z_{x_6} + .433z_{x_7} + .475z_{x_8} + .364z_{x_9}$$

Where the z's are, the variable scores expressed as standard scores.

The size and the sign of these weights must be taken into account in examining the nature of the canonical variates. For example, a positive sign for a weight means that a high standard score for its variable would make a positive contribution to the size of the canonical variate and a low score would make a negative contribution. The converse is also true. Therefore, the size of these weights is an index of the relative importance of each variable; whereas, the signs indicate the nature of the contribution made by the variables. For the numeration unit, for example, mental age, reading score, attention score, question$_1$ and question$_2$ made positive contributions to the pupil's rate of learning while pupil's verbal IQ and the total number of skills worked during the previous school year contribute negatively to the pupil's learning rate. Mental age seemed to be the largest contribution to the first canonical R for the numeration unit.

The interpretation of the betas has the same problems as are found in the case of multiple regression analysis in that they are typically quite unstable from sample to sample particularly where relatively small samples are used. However, the interpretation of a given canonical correlation is greatly aided when the structure R's (factor structure loadings) are used to examine the relative contributions made by the individual variables. The structure R indicates the strength of relationship (correlation) between the original variable and the derived canonical variate (factor); hence, it does not change as radically from sample to sample as the betas do. Therefore, it seemed to be more appropriate and meaningful to use the structure R's to examine the strength of association between the predictor variables and rate of learning. Table 10 lists the structure R's for the first canonical variates of each unit selected for this study.

Total number of skills completed in 1967 seemed to relate to the derived first canonical variate consistently higher than other variables for all six units (Table 10). Hence, it may be interpreted as the best predictor for the criterion domain—the rate measures.

Table 11 presents the proportions of the variance of the predictor set extracted by the first canonical variate (factor) for each of the selected units of IPI math curriculum; and Table 12 presents the proportions of the variance of the criterion set extracted by first canonical variate for each of the units.

The first column of Tables 11 and 12 lists the first canonical correlation of each unit; the second column lists the R^2 (λ) of the first canonical

Table 10.—Correlation Between Original Variables and Derived First Canonical Variates for Each of the Selected E-Level Units (Structure R's) of the IPI Mathematics Curriculum

		Criterion Set (Rate Measure)				Predictor Set (Student Characteristics)								
Unit	Canonical R	Rate$_1$	Rate$_2$	Rate$_3$	Rate$_4$	Verb I.Q. (X_1)	Non-verb I.Q. (X_2)	M.A. (X_3)	Math Achiev. (X_4)	Read Achiev. (X_5)	Skills 1967 (X_6)	Atten. (X_7)	Ques.$_1$ (X_8)	Ques.$_2$ (X_9)
Numeration	.438**	.087	-.678	-.053	.669	.152	.320	.380	.310	-.313	-.260	.309	.531	.469
Place Value	.485*	.849	.818	.486	-.104	.024	-.053	.594	.682	.473	.684	-.180	-.041	-.256
Addition	.689	-.866	-.689	-.449	-.080	-.168	.191	-.279	-.233	-.420	-.658	-.572	-.121	-.256
Subtraction	.656**	.780	.672	.690	.361	.014	.315	.299	.342	.426	.554	.301	.372	-.072
Multiplication	.473*	-.466	-.588	-.851	-.316	-.220	-.166	-.265	-.607	-.097	-.724	-.355	-.120	-.010
Comb. of Processes	.529	.128	.068	-.767	.754	-.225	.147	.282	.087	.161	.720	.166	.321	.276

**.01 level of significance
*.05 level of significance

Table 11.—The Components of Redundancy Measures for the First Canonical Vector of the Predictor Set (Student Characteristics) for Each of the Six Mathematics Units Indicated for this Study

Unit	Canonical R	R-Squared λ	Variance Extracted v_p	Redundancy $\lambda \cdot v_p$	Proportion of Total Redundancy
Numeration	.438**	.192	.125	.024	.585
Place Value	.486*	.236	.174	.042	.532
Addition	.689	.476	.135	.064	.609
Subtraction	.656**	.431	.116	.050	.561
Multiplication	.747*	.224	.132	.030	.612
Com. of Processes	.529	.280	.101	.028	.341

** .01 level of significance
* .05 level of significance

R which represents the proportion of the variance of the composites shared by the first canonical vectors of the two sets of variables. Column three of the tables presents the proportion of the variance of the total set extracted by the first canonical variate. The fourth column lists the amount of redundant variance attributed to the first canonical factor and the last column expresses the values in the fourth column as proportions of the total redundancy. The redundancy is obtained by finding the product of R-squared (λ) and the variance extracted and, hence, tells us what proportion of the variance extracted from the total predictor set by the first canonical vector of that set is associated with the first canonical vector in the criterion set. It will be noted that these proportions—the redundancy—are very small for all six units. However, the proportions in the last column, showing what proportion the redundancy for the first canonical vector is of the total redundancy for all canonical vectors of the predictor set, indicate that in five of the six units the first canonical variate accounts for over 50 percent of the total associated variance.

As an example in interpreting Table 11, the first canonical R for the numeration unit is .438, which is significant at .01 level, and $\lambda = .192$ which means that 19 percent of the variance in the first canonical variate of the predictor set is predictable from the first canonical variate of the criterion set. The proportion of the variance of the predictor set extracted by the first canonical

Table 12.—The Components of Redundancy for the First Canonical Vector of the Criterion Set (Rate Measures) for the Six Mathematics Units Indicated in this Study

Unit	Canonical R	R-Squared λ	Variance Extracted v_c	Redundancy $\lambda \cdot v_c$	Proportion of Total Redundancy
Numeration	.438	.192	.230	.004	.494
Place Value	.486	.236	.409	.096	.701
Addition	.689	.476	.358	.170	.632
Subtraction	.656	.431	.416	.179	.765
Multiplication	.474	.224	.347	.078	.619
Comb. of Processes	.529	.280	.295	.083	.519

variate is .125. This value is obtained from the column sum of the squared loadings (squared structure R) of variables within the predictor set on the first canonical variate divided by the number of variables. Therefore, it indicates the variance extracted by the first canonical variate of the predictor set for the numeration unit is .024. This value is obtained by multiplying the λ and the proportion of variance extracted by the first canonical factor (columns 2 and 3 Table 11). Therefore, it indicates the amount of redundant variance attributed to the first canonical variate of the predictor set. The last column of Table 11 indicates that, for the numeration unit, 59 percent of the total variance of the predictor set predicted by the variance in the criterion set is associated with the first canonical variate. This value is obtained by dividing the redundancy of the first canonical factor (column 4 of Table 11) by the sum of the redundant variance of all the possible canonical factors—four in this case.

The data presented in Tables 11 and 12 seem to indicate that the components of the redundancy measure of the first canonical factor is quite consistent throughout the six sample groups (six units). In the predictor set, (Table 11) the proportion of the total redundancy associated with the first canonical variate is very similar except for the combination of processes unit. Although the components of the six units for the criterion set are not as consistent as the predictor set (Table 12), the differences are not great. Therefore, in general, it can be concluded that there is a relationship between the two domains of measures being studied and that the percentages of associated variance related to the first canonical R are quite similar for the six sample groups included in this study.

The overall results of the study on pupil's rate of learning in IPI mathematics in the elementary grades seemed to substantiate the hypothesis that student variables, such as the verbal and nonverbal IQ, reading achievement, mathematic achievement, achievement in IPI mathematics during the previous school year, student attention as well as his reactions to the difficulty level of the work, and his reactions to the learning material used are related to the individual pupil's school achievement. Through the use of the canonical correlation method, we were able to test the general hypothesis that these selected predictor variables are related to how fast a particular student learns the material under an individualized instruction program.

However, because of the fluctuation of the relative importance of the predictors from unit to unit as indicated by the results of Multiple Correlation analysis, as well as the canonical correlation analysis, the results of the analyses do not offer a clear cut explanation about the nature of the relative contributions made by each of the predictors to the predictions of the rate measures. It may be concluded that rate of learning is specific to a given task and is not a general factor characterizing student performance in all learning situations. Therefore, it is the writer's judgment that, before any consistent relationship between the rate measures and the predictors can be established, a careful study of the characteristic nature of the task to be learned should be carried out. A logical extension of this study would be to further examine the nature of the learning objectives and the types of skills to be learned within each of the units of the IPI curriculum, so that relationship between pupil's rate of learning and the type of tasks to be learned can be specified.

FOOTNOTE

1. The research reported herein was supported by the Learning Research and Development Center supported as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education.

REFERENCES

1. Bolvin, John O., report on Individualized Instruction Project presented at the Board of Visitors Meeting, Learning Research and Development Center, University of Pittsburgh, April, 1965, (Mimeographed).
2. Carroll, John B., "A Model of School Learning," *Teachers College Record*, 64:723–732, 1963.
3. Cooley, William W.; Lohnes, P. R., *Multivariate Procedures for the Behavioral Science*, John Wiley, New York, 1962.
4. Cooley, William W.; Miller, Judy D., "The Project TALENT Test as a National Standard," *Personnel and Guidance Journal*, 44:1038–1044, June, 1965.
5. Guilford, J. P., "Three Faces of Intellect," *American Psychologist*, 14:569–679, 1959.
6. Hotelling, H., "The Most Predictable Criterion," *Journal of Educational Psychology*, 26:139–192, 1935.
7. Kendall, M. G., *A Course in Multivariate Analysis*, Hafner, New York, 1961.
8. Lindvall, C. M.; Bolvin, John, "Program Instruction in the Schools: An Application of Programming Principles in Individually Prescribed Instruction," *Program Instructions, 66 Year Book of National Society for the Study of Education, Part II*, University of Chicago Press, Chicago, 1967.
9. Sjogren, Douglas D., "Achievement as a Function of Study Time," *American Educational Research Journal*, 4:337–343, 1967.
10. Suppes, Patrick, "Modern Learning Theory and Elementary School Curriculum," *American Educational Research Journal*, 1:79–94, 1964.
11. Yeager, John L., "Measures of Learning Rates for Elementary School Students in Mathematics and Reading Under a Program of Individually Prescribed Instruction," unpublished doctoral dissertation, University of Pittsburgh, 1966.
12. Yeager, John L.; Lindvall, C. M., "An Exploratory Investigation of Selected Measures of Rate of Learning," *Journal of Experimental Education*, 36:78–81, Winter, 1967.
13. Wang, Margaret C., "An Investigation of Selected Procedures for Measuring and Predicting Rate of Learning in Classrooms Operating Under a Program of Individualized Instruction," unpublished doctoral dissertation, University of Pittsburgh, 1968.

FIELD NEWS

Willard E. North

Central Missouri State College

THE AMERICAN ASSOCIATION FOR JEWISH EDUcation has published a report on "The Teaching of Jewish Civics in Jewish Schools in the U. S." Jewish Civics includes the study of the structure, needs, and problems of the Jewish community. The survey, published as monograph No. 33, involved the sampling of 820 orthodox, conservative, and reformed Jewish schools enrolling over 219,000 pupils.

The study dealt with the organization of the Jewish Civic program, various components of the curriculum, instructional methods, and methodology used. The status report on the program is supplemented by a series of recommendations designed to improve its effectiveness. This study was conducted by Dr. Gerhard Lang. Copies of the report may be obtained from the Association for Jewish Education, 101 Fifth Avenue, New York, New York 10003. Cost: $1.25 per copy, postpaid.

* * * * * *

MOST OF THE NATION'S COLLEGES AND UNIVERsities are not currently engaged in programs designed to assist disadvantaged students. This was the conclusion reached in the first phase of a comprehensive study of Collegiate Compensatory Programs for Disadvantaged Youths being conducted at Teachers' College, Columbia University. This study is under the direction of Professor Edmund Gordon, Chairman of the Department of Guidance at Teachers' College. Sponsored by the Ford Foundation in cooperation with the College Entrance Examination Board and the National Scholarship Service and Fund for Negro Students, the study involves a large sample of colleges and universities throughout the country. The objective of the study is to determine the number of institutions with special programs, and the nature of these programs which are designed to make college experience more accessible and successful for disadvantaged college aspirants. The study includes universities, 4-year colleges, institutions offering programs of less than 4 years, and nursing schools. Later stages of the study will attempt to determine the effectiveness of these programs.

Seventy-two percent of the institutions responding stated that they currently did not have such programs in existence. The most frequent explanation given was that the colleges did not perceive themselves as realistically being able to initiate special compensatory programs as a consequence of their institutional structure and purpose. Slightly more than a tenth of the institutions that did not have compensatory programs had formulated plans to begin them. Less than 1 percent of the institutions responded that they once had such programs but had already discontinued them.

The group of institutions which reported the highest percentage of programs for disadvantaged students were the universities (56.7 percent). Thirty-two percent of the community colleges had such programs while 30 percent of the 4-year colleges did. However, according to Dr. Gordon, this comparison may be misleading because many of the community colleges feel that their total program reflects the need for a non-typical college-bound population and are therefore not regarded as special programs. Only 5 percent of the nursing schools responded that they had programs for disadvantaged students. A casebook containing studies of selected programs for disadvantaged students is expected to be published at the conclusion of this study.

* * * * * *

INSTRUCTIONAL AND PRACTICE IN SYSTEMATIC observation is an important part of five short lessons on evaluation of human learning behavior that has been prepared by Dr. Stephen Yelon, an educational psychologist, at Michigan State University. The lesson units, each less than an hour in length, consist of programmed audio–visual materials (although the trainees themselves use paper and pencil). Materials are intended for use in learning carrels, where each trainee may work at his own pace.

In the units, movie and slide film presentations are used to simulate real-live observational situations. Dr. Yelon believes that this simulation has several advantages over asking students to practice in live situations. Essentially, the advantages lie in the control that can be exercised over what the student is to observe. For instance, simulation offers the possibility of easily moving the student from simple to more complex situations, whereas the complexity that a live situation may take on is almost impossible to predict. Then too, simulated situations provide an advantage over sending students on field trips to observe, for with simulation the instructor himself knows what the student has seen and can judge how well the observations have been made. Also, unlike reality, the simulated situation can be replayed until the lessons the instructor believes are present have been learned. In the units, trainees are taught two principal observational lessons. One is to make relevant observations—to observe and record only what is genuinely relevant to the observation problem rather than to be distracted by peripheral occurrences. The second is to make observations in operational terms, that is, to make reports that are actual records of behavior or circumstances rather than judgments or impressions.

Practice in making observations, via the simulated situations, is included. When the trainees have made their observations, they immediately compare their record with "good" observations included in the programmed material. Dr. Yelon perceives his preparation of the program units as having two objectives. One is the development of individualized programs of observation training. The other is the development of an instrument by which the feasibility of simulation for mass use in teacher training can be tested. In conjunction with the latter purpose, he plans group testing of the 5-unit series. Questions regarding the observation training units may be addressed to Dr. Stephen Yelon, Educational Development Program, Michigan State University, East Lansing, Michigan 48823.

* * * * * *

FOR NEARLY A DECADE, COLLEGE BOUND HIGH school students have been urged to apply for admission to a number of colleges—at least three, and as many as six. A study by the Regional Education Laboratory for the Carolinas and Virginia (RELCV) indicates that more than half of the 1967 Virginia high school graduates applying to colleges filed only one or two applications. More than 86 percent of these students, the study indicates, were accepted by their first- or second-choice college.

"Multiple College Applications" by Andrew Bean, cites research conducted in 1968 among Ivy League schools which showed that 62 percent of the 78,000 students applying to these colleges filed only one admissions application. Similarly, a 1966 study in California, Illinois, Massachusetts, and North Carolina revealed that half the students applying for admission filed only one application. The RELCV study conducted in cooperation with the Virginia State Board of Education concludes that "it is the brighter student applying to the more selective colleges who files the most applications. This contradicts the popular belief that it is the student who is less talented academically who files the greatest number of applications."

Other conclusions of the study include the following: (1) College admission officers tend to give high school class rank twice as much weight as College Entrance Examination Board test scores, but students with high test scores and a low class ranking are the ones most likely to file more than three admission applications. (2) Thirty-five percent of the Virginia students surveyed applied to only one college, and 85 percent of these were accepted for admission. Only 3 percent filed six or more applications. (3) Seventy-seven percent of the students applying to only 2-year colleges filed just one application. (4) Students should apply to colleges that have previously accepted the student with test scores and high school class ranking similar to theirs. These are the colleges most likely to accept them. The Multiple Admissions Application problem, the study concludes, is more a myth than a reality. In Virginia, less than 16 percent of the high school students going on to college filed three or more applications. Nationwide, more than half the high school seniors going to college decide on one college and make their application there.

The RELCV report calls for an evaluation of possible cooperative admission activities among groups of colleges which tend to draw their students from the same pool of applicants. A cooperative admissions system, the study suggests, might ease these problems. It could employ a single application form for the cooperating colleges and use a centralized system for preparing application forms for admission decisions.

* * * * * *

"A LONGITUDINAL STUDY OF THE BARRIERS affecting a pursuit of Higher Education by New York High School Seniors—Phase I" has been published by the Bureau of Research in Higher and Professional Education of the New York State Education Department. The objective of this first phase of the study was to scrutinize the restrictions which 5,175 high school seniors in New York State came upon as they attempted to secure education beyond high school. Nearly two-thirds of the participating students indicated plans to continue their education. Of those students who intended to continue, the evident trend was to select publicly supported institutions. Twenty-nine percent of the students studied submitted applications to the 4-year units of SUNY. Those students who decided to submit applications to private schools were likely to apply to institutions outside of New York State.

A number of "intervening" variables in the college decision were considered. It was indicated that girls in the upper quartile of their graduating classes were less likely to attend college than boys from the same quartile. The students' socioeconomic class was an important variable, in that the higher the class the greater the probability that the college application would be filed. Proximity to an institution of higher education did appear to boost the college-going rate. Only 3 percent of seniors studied said they did not live within 25 miles of an institution of higher education. It was noted that the individual counseling received by the college bound seniors was greater than the time spent with the noncollege group.

Another conclusion indicated that parents exerted a great deal of influence upon college-bound students, while the non-college-bound sought counsel from individuals outside the educational establishment and reported infrequent communication with adults on future plans. Parents were somewhat more reluctant to assume financial obligations for their daughters' future education than for their sons'. Over half of the total group requested financial assistance during the freshman year. The financial barrier did not seem to prevent enrollment plans but may be more closely related to college attrition rather than to the actual decision to go to college. In January 1969 a follow-up study of 20 percent of the population was conducted. These findings will be released in a separate publication. However, copies of the report of Phase I of the study are available from the Bureau of Research in Higher and Professional Education, Albany, New York 12224.

* * * * * *

A SENIOR VISITING FELLOWSHIP TO SUPPORT research into the relations between humanities and social science and technology for outstanding scholars has been established at Cornell University. Research is expected to include such topics as the investigation of the influence of scientific conception of the universe upon literature, historical analysis of social resistance to technological change, and the use of scientific technology in humanistic studies.

Available for the first time during the 1970–71 academic year, the fellowship is sponsored jointly by the University's Society for Humanities and its Program on Science, Technology and Society. The fellow will be selected among candidates from all over the world, according to Max Black, Director of the Society for the Humanities. Applicants, he said, must have considerable expertise in a humanistic discipline, in pure or applied science, in technology, or in any combination of these. The fellow will be expected to engage in original research on some well-defined topic under the general heading of "Relations Between the Humanities, Science and Technology." In addition, he will be expected to engage in a limited amount of informal teaching related to his research, in ways that might be helpful to the Program on Science, Technology and Society.

The fellow will be in residence during the academic year and will be provided with a study and secretarial assistance and will have full access to the libraries and other resources of Cornell University. Further information can be obtained by writing Max Black, Director for the Society of Humanities, Cornell University, Ithaca, New York 14850.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

THOMAS ARENA received from Ohio University the degrees BS in elementary education, MEd in school administration, and the PhD in elementary education, specializing in reading and language arts. He is presently associate professor of education at Slippery Rock State College, Slippery Rock, Pennsylvania. Previous positions include those of elementary teacher, as well as supervisor of elementary education and coordinator of federal programs at the county level. He has also served on the faculty of Ohio University as coordinator of the Teacher Corps program.

CONSTANCE T. FISCHER holds a BA in government and a teaching certificate in social studies from the University of Oklahoma. Continuing interest in the role of social values in education led to the present article, which is based on the author's master's thesis written for the Psychology Department of the University of Kentucky. She received her PhD from the same university, and completed her internship in the Veterans Administration clinical psychology training program. She is now an assistant professor at Duquesne University, Pittsburgh, where the entire psychology department is attempting to develop a rigorous philosophical foundation and methodology for psychology as a *human* science. These efforts led to the self-criticisms of the present article.

LEO R. HILFIKER's undergraduate work was completed at Wisconsin State University, Whitewater, followed by MS and PhD work at the UW Madison campus. He taught high school, and was a school district administrator. He then joined the State Department of Public Instruction holding several positions. At the time of the acceptance of this article, he was working full time with the Research and Development Center, The University of Wisconsin as Coordinator of the Planned Educational Change Project. Dr. Hilfiker was killed on February 24, 1970 in an automobile accident while traveling in the state of Wisconsin.

FRANK B. MURRAY is an associate professor in the College of Education at the University of Delaware. Previously he held the same rank in the Department of Educational Psychology at the University of Minnesota. His PhD is from the Johns Hopkins University, and his principal area of research has been intellectual development and the psychology of curriculum.

BARTON B. PROGER received his EdD in research methodology and applied statistics from Lehigh University. He also holds an MA in secondary mathematics education from Lehigh and a BS in mathematics and physics from Albright College. He is presently statistician and behavioral research scientist for Research and Information Services for Education (RISE). **RAYMOND G. TAYLOR, JR.** holds an MS and EdD from the University of Pennsylvania, a BD from Episcopal Theological School, and a BS in mathematics from Bucknell University. Dr. Taylor serves as a consultant to RISE and is presently educational director for Vanguard Schools, Haverford, Pennsylvania. **LESTER MANN** received his PhD in clinical and experimental psychology from the University of North Carolina, and is presently supervisor of special education for Montgomery County, Pennsylvania. He also serves as a co-editor for the *Journal of Special Education.* **JOHN M. COULSON** received his EdD in administration from the University of Pennsylvania. He is currently an associate director of RISE and also lectures at Villanova University. **ROBERT J. BAYUK, JR.**, is a doctoral student in educational research at the University of Pennsylvania.

CLAY SCHOENFELD is the editor of *Environmental Education*, a new journal on research and development in conservation communications. He is Professor of Journalism and Wildlife Ecology at The University of Wisconsin, Madison, Chairman of the Center for Environmental Communications and Education Studies, and Director of the UW Summer Sessions. He has served as a communications consultant to the Ford Foundation, the Carnegie Corporation, and the Continental Army Command as well as the Conservation Foundation. Prof. Schoenfeld is a charter member of the board of directors of the Gordon MacQuarrie Foundation for conservation journalism, and a member of the board of directors of Dembar Educational Research Services, Inc. He is moderator of Wisconsin's inter-college course, "Environmental Resource Management Principles and Problems."

MARGARET C. WANG earned her BS in elementary education at Mount Mercy College, her MS in early childhood education at St. John's University, and her PhD in child development and educational research at the University of Pittsburgh. Dr. Wang is presently an assistant professor of child development in the Department of Psychiatry at the Medical School of the University of Pittsburgh and a research associate at the Learning Research and Development Center of the University of Pittsburgh, serving as director of testing and evaluation for the Primary Education Project of the Center.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor Curriculum and Instruction, The University of Wisconsin, Madison.

PHILIP LAMBERT, Professor of Educational Psychology and Administration, The School of Education, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1970

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Assistant Professor of Special Education, Yeshiva University, New York.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English and Education, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLASS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

CARLETON W. WASHBURNE, Director Emeritus of Teacher Education and the Division of Graduate Studies, Brooklyn College, 2248 Kent Street, Okemos, Michigan 48864.

50th Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **OCTOBER 1970** **NUMBER 2**

ARTICLES

What Does Research Say About the Time to Begin Reading Instruction?—*Dolores Durkin* 51

A Test of Hypothesized Relationship Between Learning Time and Amount Learned in School Learning—*Ronald P. Carver* 57

Effect of Encouragement and Reassurance on Students' Performance in Remedial English—*Harold N. Beyer, Eugene R. Oetting* 59

Using Students' Notes to Examine the Role of the Individual Learner in Acquiring Meaningful Subject Matter—*Michael J. A. Howe* 61

Sensory-Feedback Modes as Determinants of Learning and Memory—*Harvey M. Sussman, Karl U. Smith* 64

A Note on the Validity of the Minnesota Teacher Attitude Inventory (MTAI)—*Erness B. Brody* 67

Mental Ability, School Achievement, and Language Arts Achievement in the Prediction of Delinquency—*Betty Ruth Raygor* 68

The Superintendent's Negotiation Role—*William E. Caldwell* 73

An Instrument for Evaluating Survey Research—*Richard L. Kohr, Marilyn N. Suydam* 78

Taped Instruction Effects on Problem-Solving Skills of Seventh-Grade Children—*Francis Sekyra III, M. Ray Loree* 86

DEPARTMENTS

Among Those Present 96

Book Reviews—*Robert E. Clasen* 58

Field News—*Willard E. North* 94

Forthcoming Articles 72

It Says in the JER 50

New Books from the Publishers 63

(The Journal of Educational Research is indexed in Abstr.S.W., CIJE, CSPA, Current Contents, Ed. Adm. Abst., Educ. Ind., M.H.R.Ind.; PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1970 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by **DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.** **WALTER FRAUTSCHI**, President. Prof. **WILSON B. THIEDE**, Vice President and Publisher. Prof. **CLARENCE A. SCHOENFELD**, Assistant to the Publisher. **ARNOLD CAUCUTT**, Treasurer and Business Manager. **NORMA COURNOW CAMP**, Supervisor of Editorial Services.

It Says in The JER...

WHEN IS A CHILD ready to learn to read? Durkin discusses the history of and implications for future research on reading readiness. She suggests more complex research on the subject to avoid beginning with a wrong question and ending with wrong and meaningless answers, of which she says, "We have enough . . . already" (p. 52).

LEARNING TIME IS an important part of classroom learning. The degree of learning, when other things are equal, is a result of time the pupil spends actively in learning, says Carver (p. 57).

SIMPLE REASSURANCE AND ENCOURAGEMENT, instead of creating a self-fulfilling prophesy, might have served in some way to reduce chances of success for students in remedial English classes, suggest Beyer and Oetting. Counselors who use simple reassurance should realize that the outcome might not be in the direction they anticipate (p. 59).

NOTES MADE BY ADULT LEARNERS provide a useful indication of the products of individual encoding processes in meaningful verbal learning and memory, says Howe. He also found that there is a significant positive correlation between the "efficiency" of note-taking and the success of subsequent recall attempts, despite the presumed individual difference in what constitutes "efficient" note-taking (p. 61).

PURELY VISUAL OR PERCEPTUAL MODE of practice and learning was just as good as active modes and significantly better than a compensatory or negative-feedback mode of tracking, Sussman and Smith suggest. They believe compensatory negative feedback control may represent the poorest mode of sensory control for memory (p. 64).

HIGH SCHOOL STUDENTS RATED their student teachers in Brody's study. Teachers scoring high on The Minnesota Teacher Attitude Inventory were also those who were rated favorably by non-college-bound students. Those scoring low were rated more favorably by college-bound students (p. 67).

THE CONTENTION THAT THERE is a negative relationship between delinquency and high school rank for ninth graders is supported by Raygor's study. She also finds some evidence to indicate that low school achievement is related to personality disorders rather than to delinquency, and that disturbed students, whether delinquent or not, show many similarities (p. 68).

TEACHERS ARE MORE SATISFIED with a non-participant superintendent negotiation role, while school board members are more satisfied with an advisory superintendent negotiation role. Superintendents in large school districts, Caldwell also found, have more authority in negotiations than superintendents in small districts (p. 73).

AN INSTRUMENT FOR EVALUATING Research Reports was developed by Kohr and Suydam and is suggested for use not only in evaluation but also in planning and developing research studies, evaluating studies for publication, comparing usefulness of conflicting results, for abstracts, and for research classes (p. 78).

SEVENTH GRADERS CAN improve problem-solving abilities through practice with exposure to taped programs. Sekyra and Loree found that increasing amounts of practice yielded significant incremental improvement on each of the four problem-solving variables investigated (p. 86).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

Diary No. 194
Date 27/1/71
File No. Lie
Bureau Ednl. Psy. Research

Editorial/Abstract

What Does Research Say About the Time to Begin Reading Instruction?

◆ This paper which examines historically the concept of "reading readiness" is most timely today when preschool and earlier school programs are increasing in number as we try to find ways to rectify the school failure of so many children of the urban poor. The influence of the recapitulation theory on psychologists and the resulting emphasis on maturation gave rise to the concept of "reading readiness" and postponement of the starting time as a solution to difficulties in beginning to read. Tests were devised to help the school determine when the child was ready, i.e., had matured sufficiently.

The 1960's brought a change with increasing attention to the importance of learning opportunities in preparation for school. A serious deficiency in readiness research is the tendency to explain correlational data in causal relationships and the failure to examine readiness and reading longitudially. Furthermore, the importance of type and quality of instruction including the materials to be used is critical as is the fact that learning to read progresses one step at a time. Success with the first step often prepares the learner to be ready for the second.

Further research needs to identify the learning demands of each methodology available to assess the relevant abilities of each learner, and recommend ways of matching the two factors.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

What Does Research Say About the Time to Begin Reading Instruction?

DOLORES DURKIN
Professor of Elementary Education, University of Illinois

QUESTIONS about the time to begin reading instruction are invariably bound up with the concept "reading readiness." Consequently, how they are answered directly reflects how this concept is interpreted.

When the term "reading readiness" was first used in the 1920's, the common interpretation was that readiness is the product of maturation; more specifically, of a certain stage in the child's development which equips him with the requirements for success in learning to read. Because this interpretation was generally accepted over four decades and, secondly, because it directly affected the readiness research that was done during the same 40 years, an understanding of why it was both proposed and accepted is important.

Although proposed in the 1920's, the initially assigned meaning stemmed from psychological ideas that were popular at the start of the century. One way to describe them is to describe some of the beliefs of G. Stanley Hall, for it was his teachings that dominated the psychological world at that time.

One teaching was that genetic factors determine the characteristics and abilities of each individual. With this as the accepted assumption, it was natural for the early years of the century to give attention to hereditary rather than environmental factors, and to the maturation process rather than to learning and practice.

Another of Hall's influential beliefs was rooted in his acceptance of the doctrine of recapitulation, a doctrine defined briefly but clearly by Hall himself in a 1904 publication:

> *The most general formulation of all the facts of development that we yet possess is contained in the law of recapitulation. This law declares that the individual, in his development, passes through stages similar through which the race has passed, and in the same order* (14:8).

Hall's acceptance of the recapitulation theory led him to support this view of man: (a) Each individual, as he grows and develops, passes through certain stages, and (b) these stages follow each other in an inevitable, predetermined order. Because the learning process was being assigned only secondary importance, progress through these "inevitable stages" was explained with a reference to maturation.

Later, the special importance assigned to maturation was continued in the writings of Arnold Gesell, a student of Hall's who gave sustained support to his teacher's beliefs. As Gesell offered his explanation of development, he referred to processes like "intrinsic growth," "neural ripening," and "unfolding behavior" (9, 10, 11). Whatever the language, however, the contention was the same: Growth and development proceed in stages; progress from one to another is dependent upon spontaneous maturation or, put more simply, the passing of time. Such a contention met with little criticism during the 1920's, and, as a result, during this time "reading readiness" first entered the educator's professional vocabulary. This accounts not only for the original interpretation but also for the quickness with which it was accepted.

Why "reading readiness" became part of the educator's vocabulary in the 1920's can be explained with a reference to the Measurement and

Testing Movement, which permeated both education and psychology by 1920 (22). Among other things, its concern for objective measurement led to an abundance of national surveys designed to uncover what and how much children were learning in school. Common among the findings was that the rate of non-promotion for first graders was considerably higher than for children at other levels and, secondly, that inadequate achievement in reading generally was the cause of retentions (5, 6, 20). What subsequently became as frequent as the reports of the two findings was the question: *Why* are first graders not succeeding with reading?

It would seem, at least in retrospect, that a multifactor explanation would be offered that might reasonably include problems related to——too large classes, too few materials, inadequate teacher preparation, wrong methodology, unmotivated children, and so on. This, however, was not the case. Because the 1920's were permeated by the ideas of Hall and Gesell, the *one* explanation that was both proposed and accepted placed the blame on the children's lack of readiness when reading instruction began. Or, in Hall–Gesell terminology, the children were having difficulty because at the time instruction was initiated, they had not yet reached that stage of development which would allow them to be successful. The solution? Within the Hall–Gesell framework, the obvious one was to have the schools postpone instruction until the children matured and became ready for it.

Mental Age Description

Once the "doctrine of postponement" was accepted, efforts were made to define that stage of development which equips children with the prerequisites for success in reading, and to define it in such a way that it could be measured and thus identified. Because of the new availability of group intelligence tests—another product of the Measurement and Testing Movement—the earliest of these efforts sought out a mental-age level that might define it. While other researchers were involved, the one whose findings were to be quoted for an unusually long amount of time was Carleton Washburne, a well-known leader in the then popular Progressive Education Movement. In a research report written by himself and Mabel Morphett in 1931, it was proposed that a child is ready to read when he has a mental age of 6.5 years, and it is only then that he should be introduced to reading (19).

Of interest is the fact that just a few years later Arthur Gates and his associates published a series of research articles in which the reported data went contrary to the Morphett–Washburne proposal. The findings of one study, for instance, led Gates to observe:

> *Reading is begun by very different materials, methods, and general procedures, some of which a pupil can master at the mental age of five with reasonable ease, others of which would give him difficulty at the mental age of seven* (7:508).

Because conflicting findings were being reported, it was only natural to wonder why the Morphett–Washburne data from a study of one teaching method in one school system (Winnetka, Illinois) were so readily accepted as being applicable to all children. I believe a combination of factors suggests an answer. For one thing, the Morphett–Washburne proposal fit in perfectly with the temper of the times in which it was made. It gave support to the doctrine of postponement because most children entering first grade have not yet reached the mental age of 6.5 years. It also supported the notion that development proceeds in stages, and it honored the Measurement and Testing Movement by being precise and objective. Finally, any attempt to explain the unusual influence of the mental-age concept of readiness must take into account the prominence of Carleton Washburne. He was not only superintendent of the Winnetka schools—widely admired and copied in the 1930's—but also one of the most prestigious leaders of the Progressive Education Movement. As a result, what Washburne said was listened to—and not only in reading. Even earlier than 1931, for instance, he had made very specific proposals about what was to be taught in arithmetic and at which mental-age level (23).

With all these facts in mind, neither his mental-age description of reading readiness nor the influence it wielded should come as any great surprise. In fact, they simply demonstrate what still continues to be true of research: the quality and general applicability of a study are not always what determine the influence of its findings.

Reading Readiness Tests

While some researchers were pursuing a mental-age description of readiness, others were initiating studies designed to measure combinations of abilities which might also add up to readiness. These attempts to construct and evaluate reading readiness tests began in the late 1920's and have been continued right up to the present (2, 18).

Early forms of the tests, like many currently in use, typically included subtests said to evaluate vocabulary development and auditory and visual discrimination abilities. The most noticeable difference, when earlier editions are compared with the most current ones, lies in the visual discrimination subtest. Now the trend is to use letters and words; earlier, simple pictures and geometrical forms were the usual stimuli. It was, in fact, this use of stimuli other than letters and words that often resulted in low coefficients of correlation for readiness test scores and subsequent achievement in reading. Gates, this time in 1939, was one of the first to make this point when he reported on a very comprehensive study of readiness scores as predictors of success with reading (8). The general conclusion of that report was phrased this way:

> *It should be noted that among the tests of little or no predictive value are many tests and ratings widely recommended in books and articles on reading readiness testing and teaching* (8:54).

Once again, Gates and other researchers who also found a reason in their data to question the predictive value of readiness tests were generally ignored. In fact, during the 1930's, the 1940's, and even into the 1950's, the typical school practice was to use composite readiness scores, sometimes along with MA data, for making decisions about when to start teaching reading. Such a practice, it must be noted, in no way violated either the doctrine of postponement or the high esteem accorded the maturation process. Still assumed was that entering first graders are too immature to learn to read and would start the year by participating in a readiness program. It was only after such participation that administrators and teachers would begin to consider, with the help of readiness test scores, when reading instruction should begin.

The 1960's

Certainly there is no need now to explain with any detail why the 1960's are referred to as a revolutionary period for early childhood education. Suffice it to say that these years became an era in which unprecedented attention went to the young child; in particular, to the unique importance of the early years for his eventual intellectual development (3, 16). It was also an era in which the productivity of learning opportunities was stressed, with very little explicit attention going to the maturation process. Thus, at least for early childhood education, it was a time that contrasted sharply with the four prior decades.

With the sharp changes it was only natural to have questions raised about the traditional interpretation and use of the readiness concept. After all, an era which assigns critical importance to learning opportunities during the pre-first-grade period is not likely to be patient with school practices that postpone reading instruction beyond the start of the first-grade year and assume the passing of time will insure a readiness for it.

As the years have shown, the typical response to the impatience was neither complicated nor imaginative. For the most part, schools simply altered the timing of traditional practices. Readiness tests were administered earlier, often in kindergarten where readiness workbooks could be found too. In first grade, reading instruction usually was started sooner than was typical for the earlier decades—although readiness programs still could be found in some first grade classrooms, especially in school districts that had no kindergartens. In a few places the change in timing was more radical: they introduced reading in kindergarten. In other areas, however, opposition to this was as great and vocal as it had been in years gone by.

With such changes as these occurring in the 1960's, what was learned about readiness, and thus about the optimum time to begin reading instruction? Not much, and for a variety of reasons.

To begin with, the decade gave very little attention to the basic nature of readiness. In fact, some who were urging that reading be started even before the kindergarten year seemed either to ignore the factor of readiness or to take it for granted that all children are ready before the ages of 5 and 6. Those who did deal with the concept merely stressed that many aspects of readiness are learned and therefore can be taught.

Some of the learned and thus teachable aspects were like what was being evaluated in readiness tests; so, with some alterations, they continued to be used in great quantities throughout the 1960's. As was mentioned earlier, one common change in the visual discrimination subtests was the use of letters and words in place of pictures and geometrical forms. A change of more recent origin is the inclusion of a letter-naming subtest. Although this change has been explained with a reference to studies that show a significant correlation between letter-naming ability and later achievement in reading (4), it does not take into account the likelihood that the pre-reader's ability to name letters was the result of much more than something that might be labeled "readiness." More specifically, to assume that a young child's letter-naming ability in and of itself leads to success in reading is to overlook the strong likelihood that the research referred to measured an ability that was one product of an out-of-school environment that will always be contributing to the child's success in school.

This particular use of research is singled out here because it helps to underscore what has always been a serious flaw in readiness research. Here I refer to its unfortunate tendency over the years to present correlation data as if they indicated cause–effect relationships and, secondly, to omit attention to possible reasons for the reported correlations. In fact, the customary procedure has been one in which researchers collect readiness scores at one point in time and reading scores at another. Correlation coefficients for the two sets of data are then reported—and that is it.

Although, in more recent years, this type of research has decreased, it has been replaced only by equally nonproductive studies. For example, it is now common to find reports in which the researcher simply takes for granted the validity of some readiness test as he uses it to assess the value of the variable under study for preparing children to be successful with reading. Thus, in recent years readiness test scores have been used as the criterion measure for evaluating such things as the use of teacher aides, special science materials, a pre-first grade summer program, Frostig workbooks, and Delacato's crawling prescriptions (1, 13, 17, 21, 25). At other times they have been used to establish the relative value of attending kindergarten and not attending, of having mothers involved in a program and not having them, of one type of maternal behavior *vs* some other type, and so on (12, 15, 24). Underlying all the reports is the implicit assumption that readiness scores really do tell how well or poorly a child will perform when reading instruction gets underway.

While one can understand the desire of researchers to use instruments that will yield "objective" data than can be statistically analyzed—how else can one get "significant" findings?—one cannot help but be disappointed with their failure to point out the possible flaws in an instrument. In fact, if researchers continue to act as if the validity of readiness tests is unquestionable, then we can hardly expect school personnel to rise above traditional practices which should have been seriously questioned right from the start. Because the most important reason for questioning both the practices and the research on which they are based still exists, it will be dealt with now in some detail.

Wrong Question

For as long as the reading readiness concept has been with us—and this is true whether maturation or learning or some combination of the two is given credit for readiness—the assessment question posed by both educators and researchers has been, "Is the child ready?" Unfortunately, such a question is the wrong one to ask because it is incomplete. It focuses only on the child, thus omitting attention to an equally important variable; namely, the reading instruction that will be available.

Another way of making this same point is to say that the traditional way of asking the question neglects the *relational* aspect of readiness. It fails to recognize that whether or not a child is ready—that is, whether or not he will be successful in learning to read—depends not only on his own abilities but also on the kind and quality of instruction that is offered. Realistically, this means that a child might be very ready if one type of instruction can be made available, but unready if another is offered.

Since different kinds of instruction make different demands of the learner, this relational dimension also points out that readiness should not be equated with a single collection of abilities, as has been the traditional practice. Instead, we should be thinking in terms of readiness*es* in the sense that one collection of abilities makes a child ready for one kind of instruction, while a somewhat different collection might make him ready to cope with another.

A Second Major Flaw

Still another flaw exists in the kind of thinking and questioning that have been typical when readiness is the concern. This is the failure to realize that the process of learning to read does not require children to be ready to learn everything at once. All of the pieces that comprise "reading ability" are not taught at once; therefore, they need not be learned all at once. Yet, as questions about a child's readiness to read have been considered in the past, answers seem to assume that he must be able to do everything—and right away. Such an assumption needs to be replaced by one which recognizes that a child learns to read, a step at a time; and that the important readiness requirement is that he is able to learn the first step. Fortunately, success with that first step often prepares him to be ready for the second.

Implications for Future Research

If the two flaws which have been mentioned are valid criticisms of the way educators and researchers have traditionally thought about readiness, then certain implications follow.

For example, if readiness is dependent both on the child's abilities and, as Ausubel has phrased it, on *"the demands of the learning task,"* then future research efforts ought to go in the direction of (a) assessing more successfully than has been done up to now the relevant abilities of each child; (b) identifying the possible methodologies for reading as well as the learning demands of each; and (c) helping teachers match children in terms of their abilities, with methodology in terms of what it requires of the learner. It must also be emphasized that, once these basic tasks have been done, it is only *longitudinal* studies that will be able to pass judgment on their success.

If, in what seems like a staggering task, it is also kept in mind that a child does not learn to read "all at once," then what will also come to the forefront is the realization that those who make decisions related to readiness need only be concerned with the question, "Is this child ready to read, given the fact that such-and-such will be the first learning requirement?"

Admittedly, the suggestion of a more complicated picture of readiness than has been traditionally inferred in the question, "Is the child ready?" also suggests the need for more complex research than has been undertaken up to now. However, to let a wrong question guide research is to inevitably end up with wrong and meaningless answers. And, we have enough of them already.

REFERENCES

1. Ayers, Jerry B; Mason, George E., "Differential Effects of *Science: A Process Approach* upon Change in *Metropolitan Readiness Test Scores* Among Kindergarten Children," *Reading Teacher*, 22:435–439, February, 1969.
2. Berry, Francis M., "The Baltimore Reading Readiness Test," *Childhood Education*, 3:222–223, January, 1927.
3. Bloom, Benjamin S., *Stability and Change in Human Characteristics*, John Wiley and Sons, New York, 1964.
4. Chall, Jeanne, *Learning to Read: The Great Debate*, McGraw-Hill Book Company, New York, 1967.
5. Dickson, Virgil E., *Mental Tests and the Classroom Teacher*, World Book Company, New York, 1923.
6. Editorial, "Educational News and Editorial Comment," *Elementary School Journal*, 33:641–655, May, 1933.
7. Gates, Arthur I., "The Necessary Mental Age for Beginning Reading," *Elementary School Journal*, 37:497–508, March, 1937.
8. Gates, Arthur I.; Bond, G. L.; Russell, D. H., *Methods of Determining Reading Readiness*, Bureau of Publications, Teachers College, Columbia University, New York, 1939.
9. Gesell, Arnold L., *The Mental Growth of the Preschool Child*, The Macmillan Company, New York, 1925.
10. Gesell, Arnold L., *Infancy and Human Growth*, The Macmillan Company, New York, 1928.
11. Gesell, Arnold L., *The First Five Years of Life*, Harper and Brothers, New York, 1940.
12. Gil, Mohindra, "Relationship Between Junior Kindergarten Experience and Readiness," *Ontario Journal of Educational Research*, 10:57–66, Autumn, 1967.
13. Goralski, Patricia J.; Kerl, Joyce M., "Kindergarten Teacher Aides and Reading Readiness," *Journal of Experimental Education*, 37:34–38, Winter, 1968.
14. Hall, G. Stanley, *The Psychology of Adolescence*, D. Appleton and Company, New York, 1904.
15. Hess, Robert, "Maternal Behavior and the Development of Reading Readiness in Urban Negro Children," Self and Society, *Yearbook of the Claremont Reading Conference*, 32:83–99, 1968.
16. Hunt, J. McVicker, *Intelligence and Experience*, The Ronald Press Company, New York, 1961.
17. Jacobs, James N.; Wirthlin, Lenore D.; Miller, Charles B., "A Follow-Up Evaluation of the Frostig Visual-Perceptual Training Program," *Educational Leadership*, 26:169–175, November, 1968.
18. Johnson, Roger E., "The Validity of the Clymer–Barrett Prereading Battery," *Reading Teacher*, 22:609–614, April, 1969.
19. Morphett, M. V.; Washburne, C., "When Should Children Begin to Read? *Elementary School Journal*, 31:496–508, March, 1931.
20. Reed, Mary M., *An Investigation of Practices in First Grade Admission and Promotion*, Bureau of Publications, Teachers College, Columbia University, New York, 1927.
21. Stone, Mark; Pielstick, N. L., "Effectiveness of Delacato Treatment with Kindergarten Children," *Psychology in the Schools*, 6:63–68, January, 1969.
22. Thorndike, Robert L.; Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*, John Wiley and Sons, New York, 1969.
23. Washburne, Carleton, "The Work of the Committee of Seven on Grade-Placement in Arithmetic," *Child Development and The Curriculum*, Chapter 16, Thirty-Eighth Yearbook of the National Society for the Study of Education, Part I, Public School Publishing Company, Bloomington, Illinois, 1939.
24. Willmon, Betty, "Parent Participation as a Factor in the Effectiveness of Head Start Programs," *Journal of Educational Research*, 62:406–410, May–June, 1969.
25. Wingert, Roger C., "Evaluation of a Readiness Training Program," *Reading Teacher*, 22:325–328, January, 1969.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

A Test of an Hypothesized Relationship Between Learning Time and Amount Learned in School Learning

RONALD P. CARVER
American Institutes for Research, Washington Office

ABSTRACT

It has been hypothesized that the degree of learning, other things being equal, is a simple function of the amount of time during which the pupil engages actively in learning. This hypothesis was tested with forty-eight students in an introductory psychology course. Classroom examinations were used to indicate amount learned; intelligence tests were used to indicate learning ability, and hours of study time reported by students were used to indicate learning time. When learning ability was controlled, statistically, the correlation between amount learned and learning time was statistically significant, r = .30, a result which provides support for the original hypothesis. Considering the attenuating effect of indicants, it was concluded that learning time is a highly important variable which should be taken into account in investigations of classroom learning.

CARROLL (1) has presented a model of school learning which involves five elements: (a) the quality of instruction, (b) ability to understand instruction—a combination of general and verbal intelligence in relative terms, (c) aptitude—the amount of time needed to learn the task under optimal instructional conditions, (d) perseverance—the amount of time the learner is willing to engage actively in learning, and (e) opportunity—the time allowed for learning. The last three variables are measured in units of time and for a particular learning situation the time actually spent learning is equal to the smallest of these three variables. From the model, Carroll hypothesizes "...that the degree of learning, other things being equal, is a simple function of the amount of time during which the pupil engages actively in learning.

Although the model was implicitly designed for grammer school or high school learning, it appears to fit higher education quite well. Therefore, the above hypothesis was tested in a university setting.

Method

The Ss were adult students enrolled in two introductory psychology courses (Class I, N = 27; Class II, N = 21). Each class met 100 minutes per week for 16 weeks.

The "quality of instruction" could reasonably be assumed to be equal for all students since both classes received the same lectures. Individual differences in "ability to understand instruction" (Learning Ability) were assessed by a short intelligence test, the Wonderlic Personnel Test (10 minute time limit). The actual amount of time that the student "actively engaged in learning" (Learning Time) was assessed by a self-reporting procedure for each student. During each class period students were asked to report on a slip of paper the number of hours (to the nearest half an hour) spent studying psychology during the week. The sum of the reported hours studied and the hours spent in class during the semester provided the indicant of Learning Time. The students were told by the instructor that the reports would not influence their grades and would never be discussed with them.

Individual differences in the "degree of learning" (Amount Learned) were assessed by the midterm and final examinations. Thus, the indicant of Amount Learned was the total number correct on these two objective examinations (multiple-choice).

Results and Discussion

Table 1 presents the intercorrelations among he three variables—Learning Ability, Learning Time, and Amount Learned. Notice that Learning Time did not correlate significantly with Amount Learned ($r = .07$). However, the hypothesis advanced by Carroll contained the qualifier "other things being equal." In this study, Learning Ability could be controlled statistically by a partial correlation. The correlation between Learning Time and Amount Learned with Learning Ability partialled out was .30 ($p < .05$). The ncrease in the correlation from .07 to .30 resulted mainly from the tendency for those individuals with the highest scores on Learning Ability to be lowest with respect to Learning Time ($r = -.42$).

Table 1.—Intercorrelations Among Learning Time, Learning Ability, and Amount Learned ($N = 48$)

	1	2	3	Mean	S.D.
Learning Time.... 1		−42*	.07	36.6	10.0
Learning Ability... 2			.48*	26.8	6.0
Amount Learned .. 3				77.1	14.9

*$p < .01$.

The statistically significant partial correlation between Learning Time and Amount Learned can be interpreted as supporting Carroll's hypothesis even though it was not 1.00. The following factors undoubtedly have contributed to the attenuation of this correlation: (a) restriction of range—the dropouts are most often the failing students, (b) test unreliability—both the 10 minute intelligence test and the multiple choice examination are not highly reliable, (c) study time unreliability—the self-reports could easily be falsified or erroneously recalled (d) degree of validity of indicants—the empirical variables in the study are only imperfect indicants of ability to understand instruction, actual learning time, and degree of learning.

According to Fishman and Pasanella (2) the criterion problem in the college learning situation is like the weather: everyone talks about it, but few do anything about it. The results of this study have implications for a partial solution to the criterion problem. If grades and examination scores at the college level are directly influenced by such a highly variable factor as "time actually engaged in learning," then classroom grades will always be a less than desirable criterion for many purposes. The use of self reports of study time as a co-criterion, or as a control variable, may drastically improve investigations of the relationship between learning at the college level and various intellectual and non-intellectual variables.

REFERENCES

1. Carroll, John B., "A Model of School Learning," *Teachers College Record*, 64: 723–733, 1963.
2. Fishman, J. A.; Pasanella, A., "College Admission—Selection Studies," *Review of Educational Research*, 30: 298–310, 1960.

◎◎◎◎◎◎◎◎ BOOK REVIEWS

ROBERT E. CLASEN
Book Review Editor

Research for Tomorrow's Schools

Cronbach, Lee J.; Suppes, Patrick (eds.) The Macmillan Company, New York, 1969.

The Committee on Educational Research of the National Academy of Education reports in *Research for Tomorrow's Schools.* It is an attempt, in the words of the Academy president, Ralph Tyler, "to clarify the history, goals, and current status of educational research." While perhaps doing damage to the nuances of the arguments presented, a summary of the committee's position can be stated as follows: educational research is in sad shape because of the low level scholarship demonstrated by professors of education. It is an old story. The position of the committee regarding the situation of today's schools due to the lack of research is startling and convincing. The argument that the fault lies with the professors of education is less persuasive, mostly because of the non-scholarly work reflected in the report. Since the book is a committee report, it is difficult to identify the work of specific committee members—but the pages are replete with oversimplifications, generalizations beyond the reasonableness of any data, and attributions for which no data could be brought to bear.

Often the committee falls back on that old hackneyed comparison between the medical profession and the educational profession. The comparison is familiar, and yet in the context in which it is presented here it seems to serve as an argument *against* the very position the committee is taking. In comparing the advances in medicine with those in education the authors cite the breakthroughs in biochemistry, bacteriology, and pharmacology. To extend the analogy, one should look to sociology, psychology, and perhaps anthropology for the breakthroughs in education. What have those disciplines contributed to the improvement of tomorrow's schools?

I highly recommend this book to colleagues who are interested in acquainting themselves with dissonant views of educational research.

James Raths
Professor, College of Education
University of Maryland

The Educating of Americans: A Documentary History

Calhoun, Daniel. Houghton Mifflin Company, Boston, 1969.

A WELL defined need in making available in one

(Continued on Page 85)

(Volume 64, Number 2, October 1970)

Effect of Encouragement and Reassurance on Students' Performance in Remedial English

HAROLD N. BEYER and EUGENE R. OETTING, Colorado State University

ABSTRACT

Recent studies in self-fulfilling prophesy have shown that the behavior of a person can be affected by the expectations of a significant other. This study was undertaken to investigate the effects of encouragement and simple reassurance on final course grades of students in two remedial English classes. Forty students were randomly assigned within their respective classes to experimental and control groups. Students in the experimental groups were reassured by a counselor with respect to their achievement potential in remedial English, whereas students in the control groups were merely thanked for their participation. Grades in the course of experimental and control groups were significantly different in the opposite direction from that predicted for the larger class, but were not significantly different in the smaller class. The results were interpreted as suggesting the possibility that encouragement and reassurance may have had the effect of reducing anxiety for some of those students in the larger class, leading to poorer performance. The teacher's more favorable attitude toward the smaller class may have compensated for this effect.

OVER THE past several years there has been an increasing interest in the effect of what has come to be labeled the self-fulfilling prophesy. Evidence for the operation of this phenomenon has been found in such diverse situations as: the placebo effect in medicine (7); duration of psychotherapy (2); performance of workers on the job (1, 3); number of percepts given in response to inkblots (4); number of human vs. animal percepts given in response to inkblots (5); and gains in IQ of children in a classroom situation (6). These experimental studies are typical, but not exhaustive, of the research investigating the operation of the self-fulfilling prophesy. The results of all of these studies suggest that the behavior of an individual can be significantly affected in the direction corresponding to the expectation of another with whom he is interacting, even when the expectation is itself without adequate foundation.

The high school and college counselor is frequently in the position of assessing a student's abilities and interests and of helping that student arrive at certain expectancies with regard to his ability to succeed in attaining his goals. The studies of self-fulfilling prophesy suggest that a positive attitude on the part of the counselor toward the ability of the student to attain selected goals might have an important impact on the actual success of the student in meeting these goals. If indeed certain expectancies are transmitted to and incorporated by the student, it would seem reasonable to assume that these expectancies may become, in some degree, self-fulfilling prophesies. Thus, one might inquire as to what happens to the student who is encouraged or reassured as opposed to the student who is not encouraged.

The purpose of this study was to determine the effects of simple encouragement and reassurance on the subsequent performance of college students in a remedial English course.

Procedure

The procedure involved a 10-minute individual interview with subjects about their attitudes toward English requirements. At the end of the 10-minute interview, half of the subjects (the experimental group) were given encouragement about their potential achievement in remedial English by the interviewer, previously identified by the English instructor as a skilled counselor; the remaining half were merely thanked for their participation in the project.

The subjects were thirty male and ten female college students enrolled in either of two remedial English classes taught by the same instructor. They were told that a counselor in the University Counseling Center was interested in assessing attitudes of college students toward English requirements and asked to participate by signing up for a 10-minute interview.

Students within each class were randomly assigned to experimental and control groups.

All students were interviewed within the first 3 weeks of the spring quarter, 1968. The interview involved a discussion of their past and present feelings and attitudes toward English. At the end of the interview, nineteen students received the following words of encouragement:

Judging from this interview you certainly impress me as one who has a good deal of potential which you have just begun to tap. Your apparent motivation (or, where considerable negative effect had been elicited, your need to see this through) suggests to me that you will do quite well this quarter!

The remaining twenty-one students were merely thanked for their participation in the project.

Results

Table 1 presents the differences between mean course grades of experimental and control groups in each class.

Table 1.—Mean Course Grades of Experimental and Control Groups in Two Classes

Class	Controls			Experimentals					
	N	M	σ	N	M	σ	Diff.	t	p
10:00AM	13	2.533	.915	15	1.538	1.05	-.995	-2.653	.02
12:00AM	6	2.667	.817	6	3.167	.753	.50	1.104	N.S.

Table 2.—Mean Course Grades and Final ACT Scores in Two Classes

	10:00 A.M.			12:00 A.M.			Mean		
	N	M	σ	N	M	σ	Diff.	t	p
Final Grades	28	2.071	1.067	12	2.917	.793	.846	2.756	.01
Final ACT Scores	25*	57.00	8.5	12	58.417	8.107	1.417	.490	N.S.

*ACT post-tests were not available for 3 students.

The effects of providing reassurance differed in the two classes. In the 10:00 a.m. class, which was twice as large as the 12:00 a.m. class, the control group received a significantly higher mean grade than the experimental group. In the 12:00 a.m. group, however, the mean grade of the experimental group was slightly higher than the control group mean although this difference was not significant.

These findings suggested additional t-tests comparing the final grades and final ACT scores between the two classes. These are presented in Table 2.

While there was no significant difference in the final ACT scores, the final grades in the two classes differed considerably.

Discussion

One interpretation of the results might be that the statistically significant difference between experimental and control groups found in the 10:00 a.m. group is a type I error, and no real difference exists. However, if the findings are not spurious, then reassurance may have the effect of reducing the student's chance of success. Most of the students in the remedial English classes were assigned because of inadequate skills. Many of them had experienced previous difficulty in English and had poor attitudes toward English courses. The encouragement may have had the effect of reducing anxiety about the course, and some of the students may have avoided the unpleasant tasks involved until it was too late for them to catch up with the rest of the class.

With respect to the differences between the two classes, one relevant factor may have been class size. Generally speaking, the larger the class, the less attention can be given to individualized instruction. The English instructor communicated to the experimenter prior to the data collection that her 12:00 a.m. class appeared to be far more cohesive and motivated than the 10:00 a.m. class. Where there was no significant difference between final ACT scores between the two classes, there was a highly significant difference in final grades. Having been more favorably disposed toward her 12:00 a.m. class, the instructor may have provided greater individual feedback to these students, thus compensating for any differential effect created by the interview. The more favorable attitude toward the class also may have led to the instructor's using a different basis for her final evaluation of their performance.

The usual recommendation in any experiment showing unexpected results is that it be replicated; however, the results bring up an interesting ethical consideration.

As originally conceived, this was viewed as a low risk experiment. The results now indicate that simple reassurance and encouragement, instead of creating a self-fulfilling prophesy, might have served in some way to reduce the chances of success. A replication of the experiment must now be viewed as having at least moderate risk for the Ss involved.

The results of this study should serve to warn counselors who use simple reassurance that there is at least a possibility that the outcome might not be in the direction they anticipate.

REFERENCES

1. Bavelas, A., "Personal Communication," in Rosenthal, Robert and Jacobson, Lenore (eds.), *Pygmalion in the Classroom: Teacher Expectation and Pupil Performance*, Holt, Rinehart, and Winston, Inc., New York, 1967.
2. Goldstein, A. P., "Therapist and Client Expectation of Personality Change in Psychotherapy," *Journal of Counseling Psychology*, 7:180–184, 1960.
3. Jastrow, J., *Fact and Fable in Psychology*, Houghton Mifflin, Boston and New York, 1900.
4. Marwit, S. J.; Marcia, J. E., "Tester Bias and Response to Projective Instruments," *Journal of Consulting Psychology*, 31:253–258, 1967.
5. Masling, J., "Differential Indoctrination of Examiners and Rorschach Responses," *Journal of Consulting Psychology*, 29:198–201, 1965.
6. Rosenthal, R.; Jacobson, Lenore, "Teacher's Expectancies: Determinants of Pupils' IQ Gains," *Psychological Reports*, 19:115–118, 1966.
7. Shapiro, A. K., "A Contribution to a History of the Placebo Effect," *Behavioral Science*, 5:109–135, 1960.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

Using Students' Notes to Examine the Role of the Individual Learner in Acquiring Meaningful Subject Matter[1]

MICHAEL J. A. HOWE, University of Alberta

ABSTRACT

In meaningful learning tasks the acquisition process is influenced by the way in which the individual learner interprets and encodes the material. Early attempts to recall written materials largely determine subsequent retention, even when the learner's errors are corrected. To provide an indication of how information was interpreted and encoded by learners, adult students were asked to write notes on a meaningful prose extract they heard, and they were later asked to attempt recall. Whereas the meaningful items recorded in a subject's notes had a .34 probability of recall one week later, items not recorded in notes were recalled on only .047 of occasions, suggesting that the notes learners make provide a useful indication of the products of individual encoding processes in meaningful verbal learning and memory.

PSYCHOLOGISTS have long known that what a person learns and remembers from a situation in which he is exposed to meaningful information depends to a great extent upon qualities of the individual learner (1, 3). The individual confronted with a meaningful situation interprets it in terms of his personal "frame of reference," and an encoded version of his interpretation is drawn upon for subsequent recall. Ausubel's (2) principle that "the most important single factor influencing learning is what the learner already knows" is clearly consistent with this view.

Miller (7) observed that the process by which people translate information into their "own words" plays an important role in perception and learning. Research such as Allport and Postman's study of rumor (1) indicates that the individual encoding processes may go well beyond such linguistic transformations. It seems that substantial changes in the *meaning* of materials may occur, so that the products of individual interpreting and coding functions may vary greatly between individuals.

Thus there is little doubt that what the learner himself brings to the situation has an important role in influencing the course of the acquisition process. Kay (6) described a study which examined the roles of adult learners in a 7-trial task that required recall of meaningful prose passages. The experiment continued over 7 weekly sessions, on each of which the subjects attempted recall and then listened to the correct passage. This procedure, which involves repeated presentation of the correct material, would seem to offer excellent opportunities for subjects to correct their imperfect retention, yet Kay found that there was no week-to-week improvement, and that from one session to the next the students tended to recall very accurately what they had themselves reproduced on the previous recall trial. Items reproduced on one trial, even if incorrect, were much more likely to be recalled on subsequent trials than previously non-recalled items, despite the repeated presentations of the latter as part of the correct passage. The results of a later study by Howe (5) substantially support Kay's (6) findings.

An explanation of Kay's results that is consistent with the earlier studies of the role of the individual in learning situations is that a person's reproductions draw upon a very stable encoded version of the original as interpreted by himself. This is clearly beneficial if retention is correct, but the fact that incorrect retention is equally liable to resist alteration provides evidence supporting Skinner's (8) emphasis upon the avoidance of errors at early stages in a learning task.

The apparent stability and importance for subsequent learning of the products of the original interpreting and encoding processes indicates a need for their closer examination. Just how does whatever the learner brings to the situation interact with the learning materials to determine the outcome? It would clearly be desirable to examine the relationship between the coded version and what is subsequently recalled. A useful approximation to direct observation of the productions of Ss' encoding processes might be provided by asking subjects to take notes on the meaningful prose materials to which they attend. The notes might give some indication of how the material was perceived and interpreted by each individual, and it would therefore be useful to examine the relationships between the subjects' notes and their later attempts at recall.

Method

Subjects and Materials. There were twenty-six undergraduate subjects, who participated as a single group during class time. The prose passage was a 160-word extract from a modern novel, Saul Bellow's "Henderson the Rain King," which was read at a rate of ninety words per minute. The passage contains a number of separate events and descriptions and was chosen because although highly meaningful its contents were likely to be unfamiliar to all subjects, and because there was little possibility of an S making correct guesses about the content of non-retained parts of the passage on the basis of memory for the beginning.

Procedure. The subjects listened to the recorded passage after reading the following instructions:

> *You will hear a prose passage, about 200 words long. On some later occasion I may ask you to see how much of it you can recall.*
>
> *This is what you should do: Listen to the passage, and while you are listening make notes on it, using this sheet of paper. The passage will be read at normal speed, so you will probably not be able to make a verbatim record. Try to keep your notes brief, using few words, but getting down what seem to you to be the important elements in the passage; that is, try to make a fairly accurate record of the meaning, but keep your notes brief, and don't worry about the precise wording and phraseology of the passage. Do not go on writing more than 30 seconds after the passage ends.*

The subjects were assured that their performance would be considered anonymous and confidential, and would not be used in any way for personal evaluation or grading. One week later, they were asked to attempt written recall of the prose extract.

Measurement of Meaningful Recall. The main intention of the scoring procedure was to measure the meaningful contents of notes and recall attempts. Following the procedure of a previous study by the author (5), which in turn used ideas suggested by Cofer (4) the 160-word passage was divided into twenty segments of approximately equal length. Segments were judged to be correctly reproduced if the essential content was transmitted. Written guidelines were provided, giving examples of acceptable reproductions. For example, one segment was "in came a man from the back of the house," and the scoring instructions specified that a recall attempt which provided at least the equivalent of "a man came in" would be judged correct.

The product–moment correlation between scores assigned by two judges provided a measure of the reliability of scoring. The correlation, $r = +.84$, indicates that the reliability of the meaningful recall measure was satisfactorily high for the purposes of the present study. A further indication of reliability was obtained by examining the level of agreement between judges in scoring the individual segments for each person's reproductions. On a mean of 93.5 percent of occasions (range = 75%–100%) the two judges provided the same score (reproduced/not reproduced) for meaningful reproduction of individual segments, which further suggests that measurement was adequately reliable.

Results

The mean number of segment items reproduced in the subjects' notes was 10.88 out of 20 (SD = 3.62). Mean recall was 4.08 items (SD = 2.39). The average number of items that appeared in both notes and recall attempts was 3.65, and the mean number that were reproduced in neither was 8.23. Thus the mean probability that a subject would recall an item that appeared in his written notes was .34, whereas the probability of recalling an item that was not in the notes was .047.

The mean number of words appearing in the notes that were taken was 32.81, and the mean number of words used in recall attempts was 50.73. There was a significant correlation between meaningful recall of the segment items and verbatim recall (number of words correctly recalled); $r = +.87$ ($p < .01$).

On the hypothesis that an "efficient" note-taking strategy was one in which the maximum number of units was communicated in the minimum number of words, the correlation was calculated between this ratio

$$\frac{\text{number of items reproduced in notes,}}{\text{number of words in notes}}$$

and meaningful recall. The correlation is statistically significant, $r = +.53$ ($p < .05$).

Discussion

The very low probability of a meaningful item not appearing in a subject's notes being subsequently recalled correctly suggests that the notes which are made provide a useful indication of how the materials are processed and encoded, although it is conceivable that the activity of note writing *per se* makes a contribution to later retention that is additional to the effects of the encoding processes that precede the production of notes. In fact, it is impossible to determine to what extent the relationship between the contents of a person's notes and the contents of recall indicates on the one hand that the notes made are a measure of the products of encoding processes that determine recall, and on the other hand that notetaking directly affects subsequent retention and recall.

In showing that the contents of recall attempts are largely determined by the first reproduction (which in the case of the present study takes the form of the notes) the present findings agree with those of Kay (6) and Howe (5). It is interesting to note that there is a significant positive correlation between the "efficiency" of note-taking, as defined above, and the success of subsequent recall attempts, despite the presumed individual differences in what constitutes the most efficacious notetaking strategies for individual subjects.

The results of the present experiment suggest that further and more detailed investigation making use of notes made by students as a way of attempting to examine the relationships between learning and individual coding processes would be justified. For instance, it would be interesting to explore individual differences in note-taking, and to examine the effects of different note-taking strategies upon subsequent recall of meaningful materials.

In conclusion, it appears that learning studies requiring students to make notes upon meaningful verbal subject matter may provide ways of examining the roles of individual learners in the course of the acquisition process.

FOOTNOTE

1. The research was supported by a grant from the Tufts University Faculty Research Fund, and a Summer Fellowship awarded to the author by the American College Testing Program facilitated preparation of the manuscript. The author is extremely grateful to Marianne Such and Jean Haley for their assistance.

REFERENCES

1. Allport, G. W.; Postman, L., *The Psychology of Rumor*, Holt, New York, 1947.
2. Ausubel, D. P., *Educational Psychology: a Cognitive View*, Holt, Rinehart and Winston, New York, 1968.
3. Bartlett, F. C., *Remembering*, Cambridge University Press, London, 1932.
4. Cofer, C. N., "A Comparison of Logical and Verbatim Learning of Prose Passages of Different Lengths," *American Journal of Psychology*, 54:1–21, 1941.
5. Howe, M. J. A., "Long-term Retention of Prose Material with Repeated Presentation and Recall," paper presented at Eastern Psychological Association conference, Philadelphia, April 1969.
6. Kay, H., "Learning and Retaining Verbal Material," *British Journal of Psychology*, 44:81–100, 1955.
7. Miller, G. A., "The Magical Number Seven, plus or minus Two: Some Limits on our Capacity for Processing Information," *Psychological Review*, 63:81–97, 1956.
8. Skinner, B. F., "Teaching Machines," *Science*, 128:969–977, 1958.

▶▶▶ New Books from the Publishers

Research

Combs, Arthur W. *Florida Studies in the Helping Professions* (Gainesville, Florida: University of Florida Press, 1969), 78 pp. Paper $2.00.

Cronbach, Lee J. and Patrick Suppes, (eds.). *Research for Tomorrow's Schools: Disciplined Inquiry for Education* (New York: The Macmillan Company. A publication of the National Academy of Education, 1969), xvii + 281 pp. Paper $3.95.

Department of Education and Science. *Statistics of Education Special Series No. 1 Survey of the Curriculum and Deployment of Teachers (Secondary Schools) 1965–66 Part I: Teachers* (London: Her Majesty's Stationery Office, 1968), xxix + 102 pp. Paper $4.50.

Herman, Wayne L., Jr. *Current Research in Elementary School Social Studies* (New York: The Macmillan Company, 1969), xii + 468 pp.

Oetting, Eugene R. and George C. Thornton III. *Exercises in Psychological Testing* (New York: Harper and Row Publishers, 1968), vii + 229 pp. Paper $3.95.

Roscoe, John T. *Fundamental Research Statistics for the Behavioral Sciences* (New York: Holt, Rinehart and Winston, Inc., 1969), xv + 336 pp. Cloth $9.95.

Teaching

Leacock, Eleanor Burke. *Teaching and Learning in City Schools* (New York: Basic Books, Publishers, 1969), xxii + 263 pp. Cloth $7.50.

Stauffer, Russell G. *Directing Reading Maturity as a Cognitive Process* (New York: Harper and Row Publishers, 1968), xiv + 498 pp. Cloth $11.75.

Stauffer, Russell G. *Teaching Reading as a Thinking Process* (New York: Harper and Row Publishers, 1968), xvii + 424 pp. Cloth $8.95.

General

Ash, Maurice. *Who Are the Progressives Now?* (New York: Humanities Press, Inc., 1969), x + 253 pp. Cloth $6.00.

Dale, R. R. *Mixed or Single-Sex School?* (New York: Humanities Press, Inc., 1969), xvi + 272 pp. Cloth $9.00.

Rubinstein, David and Brian Simon. *The Evolution of the Comprehensive School 1926–1966* (New York: Humanities Press, Inc., 1969), xii + 113 pp. Cloth $2.50. Paper $1.25.

Willingham, Warren W. and Nurhan Findikyan. *Patterns of Admission for Transfer Students* (New York: College Entrance Examination Board, 1969), vi + 47 pp.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

Sensory-Feedback Modes as Determinants of Learning and Memory[1]

HARVEY M. SUSSMAN and KARL U. SMITH
University of Wisconsin, Madison

ABSTRACT

The role of sensory feedback factors in determining learning and memory was investigated in terms of the effects of four different modes of movement control of sensory input on memory of stimulus patterns—i.e., (1) visual tracking of the pattern without manual control, (2) parallel pursuit tracking of the pattern with a hand-controlled visual cursor, (3) opposed pursuit tracking of the stimulus pattern with a hand-controlled visual cursor, and (4) compensatory negative feedback tracking in which the subject observed only the error of tracking the stimulus pattern. Results indicated that mode of sensory feedback control of sensed input of stimulus patterns during learning determined immediate memory, and that memory was superior for modes of feedback control in which the pattern of stimulation and of movement could be compared. The main conclusion was that the processes of active and passive learning and their role in determining learning and memory are specifiable and predictable as variable forms of movement-controlled sensory feedback.

THIS EXPERIMENT has studied learning and memory in terms of its sensory-feedback determinants. The aim of the experiment was to determine if the modes of feedback control of sensory stimuli during learning could influence the degree of recall of the space–time pattern of these stimuli. The hypothesis guiding design of the study was that learning is determined by the direct sensory-feedback effects of active movement and that the extent of learning and of memory depends on the modes of such feedback regulation of sensed input (2). In the experiment, we have measured the variations in reproduction of a space–time pattern of a perceived and tracked visual stimulus in relation to manual and visual control of this pattern and to the mode of the manual feedback control.

Method

This experiment applied tracking or experimental cybernetic methods to analysis of motor-sensory learning, using a laboratory hybrid-computer system to control the modes and conditions of sensory-feedback control under study. Figure 1 illustrates the computer system used to control manual–visual and visual modes of learning. As shown in the diagrams in the lower part of the figure, the subject's task consisted of using a strain-gauge hand control, consisting of a spring-bronze wand, to control an oscillograph feedback display. Lateral movements of the hand control caused compliant lateral movements of the second channel of the oscillograph display. This control was achieved by: (1) transducing the hand motion by means of the strain-gauge sensor, (2) amplifying this analog signal, (3) converting this analog signal to digital form by an analog–digital converter, (4) programming the digital computer for feedback control and response measurement, (5) reconverting the programmed digital signal to analog form, (6) amplifying this dc signal, and (7) using the amplified signal to actuate the feedback signal of the oscillograph.

The experiment required producing four modes or conditions of control of sensory-feedback of the stimulus wave pattern displayed on the oscillograph. As shown in Figure 1, these four modes consisted of: (a) parallel pursuit or following of the stimulus wave pattern displayed on the oscillograph; (b) opposed pursuit or following of the stimulus wave; (c) compensatory tracking of the stimulus pattern so that only the error of the tracking was displayed and observed; (d) visual observation only of the stimulus wave pattern. The hybrid-computer system was programmed to generate the wave pattern that was tracked. Thirty seconds after the tracking tasks were carried out, the subject was required to reproduce or project the stimulus wave pattern both spatially and temporally. As indicated in the upper part of Figure 2, this reproduction had to be made actively with the hand control and the reproduced wave pattern was observed on the feedback display. The stimulus wave patterns were randomized variable sine waves of the sort pictured in Figure 2, and were varied from trial to trial. All were of 15 seconds in duration.

In the pursuit tracking mode, the subject viewed the target wave form on the third oscillograph channel and simultaneously tracked its temporal–spatial patterning on the second channel. At the signal to begin the memory tracking trial, the subject reproduced the wave form he had just followed.

In the opposed tracking mode, the subject had to track the reverse of the target wave and thus executed the opposite but symmetrical wave form while watching the target wave. The subject matched the temporal rate but reversed the spa-

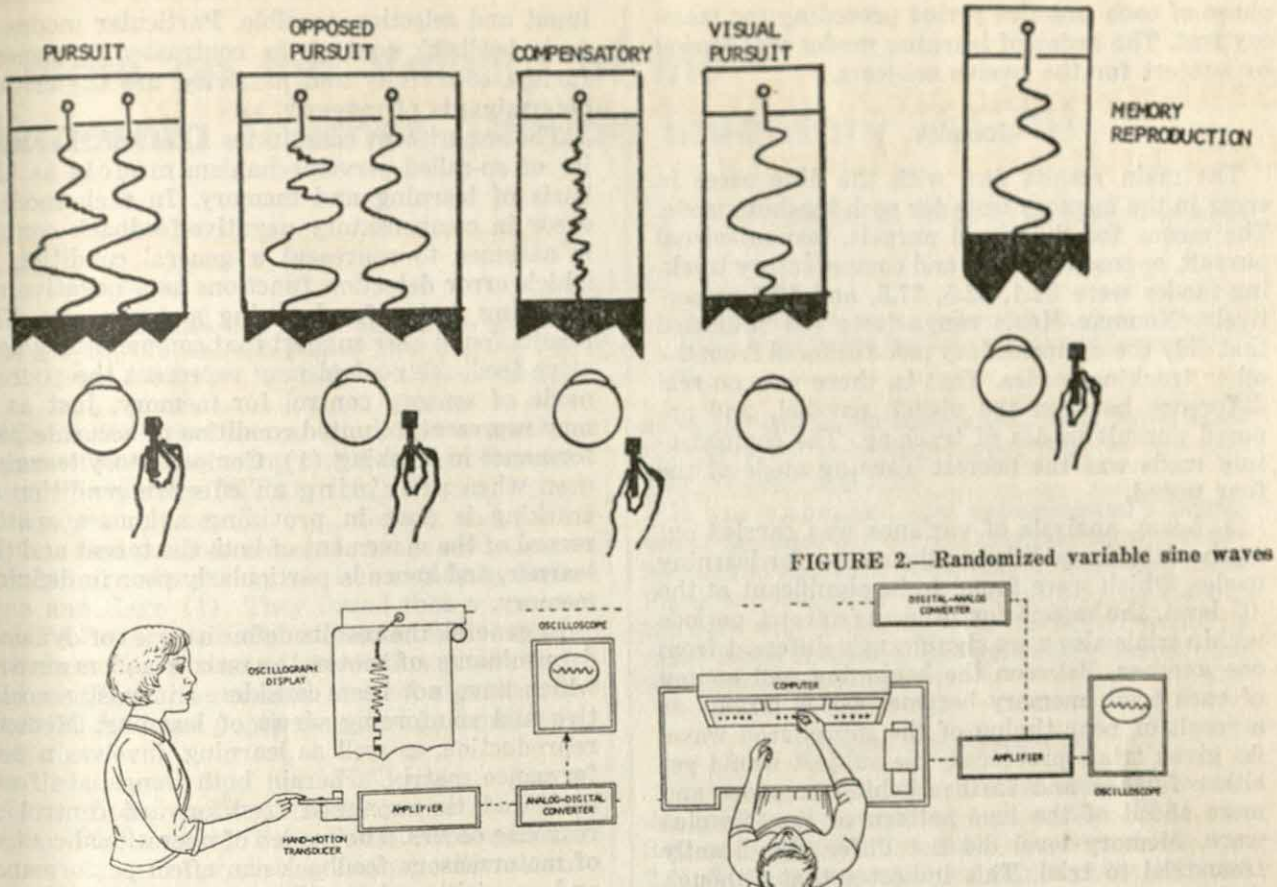


FIGURE 2.—Randomized variable sine waves

FIGURE 1.—Top: The four learning modes illustrated in terms of oscillograph records of the feedback displayer which the subjects saw and controlled. Bottom: Components of the hybrid-computer system used to control learning feedback modes, to make response measurements, and to automate different phases of the experiment by computer programming.

tial direction of the target during practice. In the memory half of the trial, the subject attempted to reproduce the actual target wave.

In the compensatory tracking mode, the subject had to position the oscillograph pen under a center null point. By compensating for the pen's movements, the subject performed the exact wave form required for the subsequent memory tracking trial. However, due to the nature of this type of tracking task, the subject never viewed the actual wave form. The operational feedback in this instance consisted of a non-informative feedback error trace of the tracking performance. The subject's only information source was the reactive feedback of his hand manipulating the transducer in the required motion paths. Again, the exact form of the tracking had to be reproduced in the memory trials.

The fourth learning mode consisted only of observing the stimulus pattern without using hand movement in the practice trials. In the memory trials, this stimulus pattern had to be reproduced by manual control as in the other memory trials.

The experimental observations were automated by means of the computer system, which was programmed for: (1) warning subjects of calibration procedures at the beginning of trials, (2) the calibration procedures, (3) warning of the beginning of a trial, (4) trial timing, (5) producing the feedback modes used, (6) varying the target wave, (7) measuring the memory of response error in memory trials, and (8) analyzing and tabulating the response measurements. In measuring the error of response in the memory trials, the computer compared the wave pattern reproduced by the hand-motion transducer with that generated during the specific learning trial that had just been tracked and measured this difference in digital voltage values. Since the sampling rate used was 128 samples per second, that number of error measurements was made each second. These measures were averaged to 5-second mean values, and at the end of the series of trials on each subject were punched out on paper tape.

The experiment used a mixed experimental design which encompassed two groups of six subjects each that differed only by sex and three repeated measures—i.e., the four learning modes previously described, five learning and related memory trials under each tracking mode, and three measurement periods within each memory trial. Only the error of performance during the memory tests was measured. The subject, however, was not informed of the fact that no measurements were made during the actual learning phase of each practice period preceding the memory test. The order of learning modes was varied by subject for the twelve subjects.

Results

The main results deal with the differences in error in the memory tests for each tracking mode. The means for the visual pursuit, manual–visual pursuit, opposed pursuit, and compensatory tracking modes were 32.1, 32.5, 37.5, and 40.7 respectively. Neuman–Keuls range tests (3) indicated that only the compensatory mode differed from the other tracking modes. That is, there was no real difference between the visual, parallel, and opposed pursuit modes of tracking. The compensatory mode was the poorest learning mode of the four tested.

A 5-way analysis of variance was carried out on the data. In addition to the means for learning modes, which were found to be significant at the .05 level, the means for measurement periods within trials also were significantly different from one another. Between the beginning and ending of each trial, memory became poorer largely as a result of poor timing of the reproduced wave. As given trials proceeded, the subject would get either further and further behind or more and more ahead of the true pattern of the stimulus wave. Memory level did not differ significantly from trial to trial. This indicates that although the amount learned may have improved from trial to trial, this improvement did not show up significantly in the dynamic memory tests used. The means for the male and female groups also were not significantly different from one another.

Discussion and Conclusions

The findings confirm the hypothesis in terms of which the experiment was designed, namely, that memory level is governed by the mode of dynamic control of sensory input during practice. The finding was that a purely visual or perceptual mode of practice and learning was just as good as the active modes, and significantly better than a compensatory or negative-feedback mode of tracking. To us there is no great mystery in this finding inasmuch as eye movements are every bit as efficient as other dynamic modes of behavior in providing dynamic feedback in tracking space–time stimulus patterns, and thus in determining their reproduction in memory. Perceptual modes of learning can be as effective for memory as other modes of learning if dynamic learning is required.

The experiment evaluates the old problems of active and passive learning in a new way by theoretically defining the factors of activity and passivity in terms of modes of movement control of sensory feedback during learning. In the feedback interpretation, there are many modes of passive and active movement control of stimulus input and selection possible. Particular modes of such feedback control, as contrasted to general factors of activity and passivity, are the critical determinants of memory.

The experiment constitutes a test of the validity of so-called servomechanism models as the basis of learning and memory. In such models, error in compensatory negative-feedback control is assumed to represent a general condition in which error detection functions as a negative reinforcing factor for learning and memory. The results found here suggest that compensatory negative feedback control may represent the poorest mode of sensory control for memory, just as it may represent a limited condition of accurate performance in tracking (1). Compensatory learning even when providing an effective condition of tracking is poor in providing a known spatial record of the movement of both the target and the learner, and hence is particularly poor in defining memory.

In general, the results define a series of dynamic determinants of motor learning and memory which have not been considered in past associative and reinforcing views of learning. Memory reproduction, as well as learning, involves a performance matrix, wherein both immediate feedback and time-spanned feed-forward control of response occurs. The modes of operational control of motorsensory feedback can affect performance and acquisition during learning, as well as performance and retention in memory. The optimal modes of feedback control in both learning and memory are those which provide maximum observable space and time compliance between movement and sensory input during learning and, in addition, make possible the accumulation of a spatial record of response in memory. Effective learning increases the spans of space and time over which the learned movement can be projected or predicted. However, in memory reproduction, the necessity for immediate feedback control of movement is not eliminated. Rather, the time spans of motorsensory feedback control can be lengthened if the task is spatially structured so that a memory record of the pattern of movement can be established and thus projected in future time.

FOOTNOTE

1. This research was supported by grants from the National Science Foundation and the Social and Rehabilitation Administration, and was carried out in the context of a training program in Psychophysiological Cybernetics of the Biological Training Section, National Institute of Mental Health.

REFERENCES

1. Poulton, E. C., "Tracking Behavior," in *Acquisition of Skill*, Bilodeau, E. (ed.), Academic Press, New York, pp. 361–410, 1966.
2. Smith, K. U.; Smith, M. F., *Cybernetic Principles of Learning and Educational Design*, Holt, Rinehart and Winston, New York, 1966.
3. Winer, B., *Statistical Principles in Experimental Design*, McGraw–Hill, New York, 1962.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

A Note on the Validity of the Minnesota Teacher Attitude Inventory (MTAI)

ERNESS B. BRODY, Rutgers University

ABSTRACT

Student teachers who scored *high* on the MTAI were rated more favorably by secondary school students who did not plan to attend college than by students who planned to attend college. The teachers who scored *low* on the MTAI were rated more favorably by their college-bound students.

RESEARCH that has attempted to predict teacher effectiveness using the Minnesota Teacher Attitude Inventory (MTAI) has not always produced consistent results, e.g., Getzels and Jackson (2). One possible reason for this lack of consistency in results is suggested in a study by Della Piana and Gage (1). They found that scores on the MTAI were positively correlated with ratings of teacher effectiveness of elementary pupils classified as *affectively* oriented and not significantly correlated for those pupils classified as *cognitively* oriented.

In the present study the MTAI was used as a measure of teacher effectiveness in a manner congruent with its use by Della Piana and Gage (1), i. e., as a measure predicting success with certain types of pupils and not with others.

Method

Subjects were sixteen male and twenty-five female student teachers attending Rutgers University. Each of these student teachers was rated by their secondary school students in two of their classes. Thus, 1,863 secondary school students also participated as subjects in the study.

Procedure

Male student teachers with MTAI scores > 19 and female student teachers with scores > 55 were assigned to an above median MTAI group ($N=18$). All other student teachers were assigned to a low MTAI group ($N=23$).

Secondary school students were classified into two groups: (1) those students who stated they planned to attend a 2- or 4-year college (college-bound) and (2) those who did not plan to attend college (non-college-bound).

Ratings of each teacher were obtained, i.e., *each* student in two of the student teacher's classes rated him on a 5-point scale (1=Excellent, 5=Poor). Each student teacher was then assigned two rating scores: first, the mean ($\bar{X}$) rating assigned to him by *all* of his students; second, a mean difference rating. This was obtained by subtracting the mean rating for non-college-bound students from the mean rating of college-bound students. Thus, negative mean difference ratings indicated that a student teacher was rated *more* favorably by his college-bound than by his non-college-bound students.

It was expected that college-bound students would be more *cognitively* oriented than non-college-bound students. Accordingly, following Della Piana and Gage (1), it was expected that student teachers who scored *high* on the MTAI would be rated more favorably by non-college-bound than by college-bound students.

Results and Discussion

The mean of the mean ratings assigned to student teachers who scored high on the MTAI was 2.62 and the mean for those who scored low was 2.51. These means were not significantly different ($t=.81$, $p<.05$).

The mean of the mean difference ratings for student teachers who scored high on the MTAI was +.12; thus indicating that these teachers received more favorable ratings from their non-college-bound students. For those teachers who scored *low* on the MTAI, the mean of the mean difference ratings was −.11. This indicates that these teachers were rated more favorably by their college-bound students. These means are significantly different ($t=2.52$, $p<.05$).

These results extend the earlier finding of Della Piana and Gage (1) to a secondary school population. Student teachers who score high on the MTAI tend to be more effective with non-college-bound students who are, presumably, *affectively* oriented. Those teachers who score low on this test tend to be more effective with college-bound students who are, presumably, more *cognitively* oriented.

These results seem to suggest the importance of considering the type of pupil with whom a particular kind of teacher might be most effective.

REFERENCES

1. Della Piana, G. M.; Gage, N. L., "Pupils' Values and Validity of the Minnesota Teacher Attitude Inventory," *Journal of Educational Psychology*, 46:167–178, 1955.
2. Getzels, J. W.; Jackson, P. W., "The Teacher's Personality and Characteristics," in Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand McNally and Co., Chicago, 1963, pp. 506–577.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

Mental Ability, School Achievement, and Language Arts Achievement in the Prediction of Delinquency

BETTY RUTH RAYGOR
Hamline University, St. Paul, Minnesota

ABSTRACT

A random sample of delinquent and non-delinquent ninth-grade boys and girls drawn from a statewide population of 11,329 were compared on high school rank, language arts achievement, and general ability with a group of "disturbed" delinquents and "disturbed" non-delinquent ninth-grade students in the state of Minnesota. The latter two groups were selected by their primed scores on scales 4 (Psychopathic deviate), 8 (Schizophrenia), and 9 (Hypomania) of the Minnesota Multiphasic Personality Inventory. These scales had been shown to be related to delinquency. The results of this study suggest that High School Rank, when combined with the "delinquency prone" rating from the MMPI, will improve the prediction of delinquency for both boys and girls, and the Cooperative English Test will enhance the prediction of delinquency for girls.

THE PREDICTION of delinquency has been a perplexing problem to individuals working in the area of delinquency. Unfortunately most studies on delinquency have been done after the fact (1, 2, 3, 4, 7). In 1947 and 1954 Professors Hathaway and Monachesi (5, 6) conducted longitudinal studies in which data was collected before the students reached the age in which delinquency rises rapidly. Using the Minnesota Multiphasic Personality Inventory, they found that scales 4 (Psychopathic deviate); 8 (Schizophrenia); and 9 (Hypomania) would increase the prediction of delinquency. Thus elevated scores 4, 8, 9 of the MMPI are sometimes referred to as the "delinquency prone" scale.

Problem and Procedure

The purpose of this study was to examine differences in measured ability, school achievement, and language arts achievement between groups of "disturbed" ("delinquent prone") students who later became delinquent and students who did not become delinquent. Both groups were compared with groups of randomly selected delinquents and non-delinquents.

This study utilized samples drawn from a larger study by Hathaway and Monachesi (6) in which 11,329 ninth-grade girls and boys from Minnesota were tested with the MMPI, and information was gathered on these subjects in relation to language arts achievement, as measured by the Cooperative English Test; general ability, as measured by the American Council on Education Psychological Examination; general school achievement, as measured by eleventh-grade High School Rank; family socioeconomic status, personality evaluation by teachers and others, police records, and other factors. These students' careers were followed up for a period of 4 to 6 years.

From this population a random sample of approximately one hundred delinquents and one hundred non-delinquents of each sex was drawn. Then this sample was returned to the original subject pool, and a sample of approximately one hundred "disturbed" (delinquent prone) delinquents and one hundred "disturbed" (delinquent prone) non-delinquents was selected. The "disturbed" category in this study constituted students who had "primed" scores (more than two standard deviations about the mean) on Scales 4, 8, or 9 of the MMPI. The "disturbed" or "delinquency prone" category is made on the basis of Professors Hathaway and Monachesi's study (5).

The "primed" score was used as Hathaway and and Monachesi (5) found that as these scores increased, the level of prediction went up. However, by using only "primed" scores, the total number of girls in the "disturbed" category was sixty-six. Thus, the study included the entire population in this instance. Only cases with complete data available were used. Invalid MMPI's were omitted from the study.

Delinquents were sorted into the four delin-

Table 1.—High School Rank Percentiles of Samples of Boys and Girls from Minnesota High Schools

	Random Delinquent		Random Non-Delinquent		Disturbed Delinquent		Disturbed Non-Delinquent	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
n	99	98	100	108	99	66	108	107
x	29.52	44.05	44.55	63.49	27.75	38.30	40.03	50.34
sd	20.80	27.40	47.07	27.44	21.63	26.63	26.72	27.22

quency categories designed by Hathaway and Monachesi (5). Only the two more serious categories were used for the boys. Three categories were used for the girls because it was felt that minor delinquent behavior on the part of girls is considered by society to be more serious than the same behavior on the part of boys, and since there were not as many delinquent girls, a more lenient definition was necessary to get an adequate sample. The sample was divided into the following classifications, (1) random delinquent, (2) random non-delinquent, (3) "disturbed" delinquent, and (4) "disturbed" non-delinquent. Comparisons using all possible combinations were made by means of the t-tests.

Results

The high school percentile ranks of boys and girls are presented in Table 1. The boys and girls in the random non-delinquent group obtained the highest mean high school rank (HSR) (boys 44.55, girls 63.49). The boys and girls in the "disturbed" non-delinquent group obtained the next highest mean HSR (boys 40.03, girls 50.34). The group of boys and girls obtaining the lowest HSR were the "disturbed" delinquent group (boys 27.75, girls 38.30).

The mean high school rank for girls is much higher than for boys. In fact, the mean HSR for random delinquent girls is almost identical to the mean HSR for random non-delinquent boys.

Table 2.—Group Comparisons of High School Rank Percentiles of Samples of Minnesota High School Boys and Girls

Groups Compared	Difference		t	
	Boys	Girls	Boys	Girls
Random non-delinquent vs. random delinquent	15.03	19.44	4.39**	5.18**
Random delinquent vs. disturbed delinquent	1.77	5.75	.58	1.33
Disturbed non-delinquent vs. random delinquent	10.51	6.29	3.14**	1.65
Random non-delinquent vs. disturbed delinquent	16.80	25.19	4.83**	5.94**
Random non-delinquent vs. disturbed non-delinquent	4.52	13.34	1.21	3.51**
Disturbed non-delinquent vs. disturbed delinquent	12.28	12.04	3.61**	2.85**

* - significant at .05
**- significant at .01

Table 3.—ACE Scores of Samples of Minnesota High School Boys and Girls

	Random Delinquent		Random Non-Delinquent		Disturbed Delinquent		Disturbed Non-Delinquent	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
n	99	98	100	108	99	66	108	107
x	87.59	90.70	92.56	96.34	86.20	88.27	83.69	87.44
sd	23.61	22.09	22.29	25.21	23.25	23.13	27.71	24.88

The results for boys and girls on HSR in the different groups are very similar. In Table 2 group comparisons on the mean high school rank, show significant differences for both boys and girls at the .01 level between the following groups: the random non-delinquent and the random delinquent; the random non-delinquent and the "disturbed" delinquent; the "disturbed" non-delinquent and the "disturbed" delinquent. There was also a significant difference at the .01 level for girls between the random non-delinquent and the "disturbed" non-delinquent and for boys between the "disturbed" non-delinquent and the random delinquents.

Thus we see that the non-delinquent in both random and "disturbed" groups had a significantly higher high school rank than any of the delinquent groups. The non-delinquent groups also showed greater variability in high school rank.

The means and standard deviations of the ACE scores for boys and girls are shown in Table 3. The random non-delinquents have the highest mean score (boys 92.56, girls 96.34) while the random delinquents have the next highest mean score (boys 87.59, girls 90.70). The "disturbed" delinquents have a mean ACE of (boys 86.20, girls 88.27) which is just slightly higher than the "disturbed" non-delinquents whose mean score is boys 83.69, girls 87.44.

The group comparisons in Table 4 show that the random non-delinquent boys and girls have significantly higher scores than the "disturbed" non-delinquents. These results are significant at the .01 level. This suggests that personality difficulties rather than delinquency is associated with lower ability.

Significance at the .05 level on the t-test was found between the random delinquents and the "disturbed" non-delinquents for both boys and girls. In comparing the boys and girls on ACE, one finds great similarity in the results. Both groups show a significant difference between random delinquents and "disturbed" non-delinquents.

Table 4.—Group Comparisons of ACE of Samples of Minnesota High School Girls and Boys

Groups Compared	Difference		t	
	Boys	Girls	Boys	Girls
Random non-delinquent vs. random delinquent	4.97	5.64	1.52	1.70
Random delinquent vs. disturbed delinquent	1.39	2.43	.42	.68
Disturbed non-delinquent vs. random delinquent	-3.90	-3.26	1.08	.98
Random non-delinquent vs. disturbed delinquent	6.38	8.07	1.97*	2.11*
Random non-delinquent vs. disturbed non-delinquent	8.87	8.90	2.53**	3.47**
Disturbed non-delinquent vs. disturbed delinquent	-2.51	-.83	.70	.22

* -significant at .05
**-significant at .01

In general, the results on the ACE are in keeping with the findings of Hathaway and Monachesi (5, 6). They indicate that there are no significant differences between delinquents and non-delinquents in ability as measured by the ACE. These results are supported by the fact that there are no significant differences among the random non-delinquent and the random delinquent. However, the present study did find a significant difference among both girls and boys when the random non-delinquents were compared with the "disturbed" delinquents.

The means and standard deviations for the Cooperative English Test Scores for boys and girls is presented in Table 5. As in HSR and ACE, the random non-delinquent has the highest mean score (boys 113.92, girls 146.81) on the Coop English Test. The random delinquent is slightly lower with a score of boys 107.01, girls 136.13. The boys mean scores for the "disturbed" non-delinquent and the "disturbed" delinquent are practically the same (106.62 and 106.14). However, in these two groups the mean scores for the girls are the reverse of the boys. The "disturbed" delinquent

Table 5.—Cooperative English Test Scores of Samplesof Girls and Boys from Minnesota High Schools

	Random Delinquent		Random Non-Delinquent		Disturbed Delinquent		Disturbed Non-Delinquent	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
n	99	98	100	108	99	66	108	107
x̄	107.01	136.13	113,92	146.81	106.14	132.74	106.62	129.84
sd	33.94	33.78	35.06	32.10	33.28	35.63	36.70	34.89

girls have a higher mean score (132.74) than the "disturbed" non-delinquents (129.84).

It is interesting to note the differences in mean scores between the boys and girls on the Coop English Test. Not only are the girls' scores much higher but the lowest mean score for girls is 129.84 ("disturbed" non-delinquent), and this score is still higher than the highest boys' score which is 113.92 (random non-delinquent). Thus, it appears that all groups of girls have higher English achievement scores than the boys regardless of the group in which the boys are placed.

Table 4 compares the various groups of boys and girls on the Coop English Test. There were no significant differences for boys. Thus, there appear to be no significant differences between delinquent and non-delinquent boys on achievement as measured by the Cooperative English Test.

The group comparisons in Table 6 on Cooperative English Test scores for girls indicate some interesting findings. As one would expect, the random non-delinquents have significantly higher scores than the random delinquents and the random non-delinquents also have higher scores than the "disturbed" delinquents. However, the random non-delinquents had significantly higher scores than the "disturbed" non-delinquents. The significant differences found above are identical to those found for girls on ACE and HSR. However, the boy's groups differ markedly, particularly on Coop English scores, where there were no significant differences between any of the boys' groups.

Conclusions

Significant differences were found, favoring the non-delinquent, for both boys and girls on High

Table 6.—Group Comparisons of English Cooperative Test Scores of Samples of Minnesota High School Girls and Boys

Groups Compared	Differences		t	
	Boys	Girls	Boys	Girls
Random non-delinquent vs. random delinquent	6.91	10.68	1.41	2.53**
Random delinquent vs. disturbed delinquent	.87	3.39	.24	.62
Disturbed non-delinquent vs. random delinquent	-.39	-6.29	.08	1.31
Random non-delinquent vs. disturbed delinquent	7.78	14.07	1.60	2.68**
Random non-delinquent vs. disturbed non-delinquent	7.30	16.97	1.46	3.70**
Disturbed non-delinquent vs. disturbed delinquent	.38	-2.90	.10	.84

* - significant at .05
**- significant at .01

School Rank (Table 1 and 2); ACE (Table 3 and 4); and Cooperative English Test (Table 5 and 6) with the exception of the Cooperative English test for boys in which no significant differences were found. When the random non-delinquent was compared with both the "disturbed" delinquent and the "disturbed" non-delinquent, a significant difference was found in the ACE for both boys and girls and significant differences in the Cooperative English Test and High School Rank for girls. It appears that the "disturbed" delinquent and the "disturbed" non-delinquent have comparable ability and in the case of girls have comparable language arts achievement and general school achievement. These results seem to indicate that the students with 4, 8, or 9 "primed" scale scores on the MMPI demonstrate many similarities regardless of whether they are delinquent or not. Although the "disturbed" non-delinquent girls have lower ability and language arts achievement scores, they have a significantly higher High School Rank than their "disturbed" delinquent counterpart. These results tend to support the contention that there appears to be a negative relationship between delinquency and High School Rank. Results suggest that even among the delinquents who remain in high school during the eleventh grade there is a negative relationship between delinquency and High School Rank. There is also some evidence that low school achievement is related to personality disorders rather than to delinquency, and that disturbed students show many similarities, whether they are delinquent or not.

REFERENCES

1. Cooper, Clara Chassell, "A Comparative Study of Delinquents and Non-Delinquents," *Psychological Service Center Journal*, 1960.
2. Dinitz, Simon and others, "Delinquency Proneness and School Achievement," *Educational Research Bulletin*, 36:131–36, April 1957.
3. Glueck, Sheldon; Glueck, Eleanor, *The Problem of Delinquency*, Houghton Mifflin Company, Boston, 1959.
4. Glueck, Sheldon; Glueck, Eleanor, *Unraveling Juvenile Delinquency*, Harvard University Press, Cambridge, 1957.
5. Hathaway, Starke; Monachesi, Elio, *Adolescent Personality and Behavior—MMPI Patterns of Normal, Delinquent, Dropout and other Outcomes*, University of Minnesota Press, Minneapolis, 1963.
6. Hathaway, Starke; Monachesi, Elio, *Analyzing and Predicting Juvenile Delinquency with the MMPI*, University of Minnesota Press, Minneapolis, 1953.
7. Kvaraceus, William C., *Juvenile Delinquency and the School*, World Book Company, New York, 1945.

Forthcoming Articles

Cohen, Daniel; Pace College, New York
"DIFFERENTIATING MOTIVATIONS UNDERLYING VOCATIONAL CHOICE"

Dacey, John S. and George F. Madaus; Boston College, Chestnut Hill, Massachusetts
"AN ANALYSIS OF TWO HYPOTHESES CONCERNING THE RELATIONSHIP BETWEEN CREATIVITY AND INTELLIGENCE"

Keenen, Charles B.; Germanna Community College, Fredericksburg, Virginia
"COMPARATIVE ACHIEVEMENT OF TWO-SUB STANDARD GROUPS OF UNIVERSITY FRESHMEN"

Mehrens, William A.; Michigan State University, East Lansing and Bruce G. Rogers; Arizona State University; Tempe
"RELATIONS BETWEEN GRADE POINT AVERAGES AND COLLEGIATE COURSE GRADE DISTRIBUTION"

Niedermeyer, Fred C.; Southwest Regional Laboratory for Educational Research and Development, Inglewood, California
"THE EFFECTS OF TRAINING ON THE INSTRUCTIONAL BEHAVIORS OF STUDENT TUTORS"

Piland, Joseph C. and Elmer A. Lemke; Illinois State University, Normal
"THE EFFECT OF ABILITY GROUPING ON CONCEPT LEARNING"

Robertson, James R. and John D. Haas; Carnegie-Mellon University, Pittsburgh
"TEACHER PERSONALITY AND THE NEW SOCIAL STUDIES"

Shaw, Ralph L. and Norman P. Uhl; Emory University, Atlanta, Georgia
"CONTROL OF REINFORCEMENT AND ACADEMIC ACHIEVEMENT"

Shively, Joe; Purdue University, Lafayette, Indiana
"CHARACTERISTICS OF STUDENTS WHO COULD NOT TAKE AND SCHOOLS WHICH DID NOT OFFER DRIVER TRAINING"

Williams, John D., Steven D. Harlow, and Del Gab; University of North Dakota, Grand Forks
"A LONGITUDINAL STUDY EXAMINING PREDICTION OF DOCTORAL SUCCESS: GRADE POINT AVERAGE AS CRITERION, OR GRADUATION VS. NON-GRADUATION AS CRITERION"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

The Superintendent's Negotiation Role

WILLIAM E. CALDWELL
The Pennsylvania State University

ABSTRACT

In an attempt to examine the role of the superintendent of schools in negotiations between teachers' organizations and boards of education, 272 school districts in twenty-six states were studied. Data collected were analyzed to test the relationships between the superintendent's negotiation role and the school district size; and the superintendent's negotiation role and superintendent, teacher, and board member satisfaction with that role. Results show that considerable conflict exists in the perception of the superintendent's negotiation role; teachers are more satisfied with a non-participant superintendent negotiation role; and board members are more satisfied with an advisory superintendent negotiation role.

THE EVER-INCREASING practice of collective negotiations between teachers' organizations and boards of education has required the superintendent to take a role in these negotiations, ranging from mere observer or non-participant to negotiator with full authority (8, 9).

The negotiation role taken by the superintendent and his behavior in that role as perceived by the teachers and board members could be influential to his success or failure as an educational leader. In an attempt to provide empirical evidence in this area, a study was undertaken to ascertain the negotiation role assigned to the superintendent and to test the relationship between the negotiation role and the satisfaction of the negotiating parties; the superintendent, the teachers, and the board members.

The study dealt with four basic questions. What relationship, if any, is there between the size of the school district and the superintendent's negotiation role? What relationship, if any, is there between the negotiation role of the superintendent and the satisfaction of (1) the superintendent, (2) the teachers, and (3) the board members.

While the superintendent can and may play a negotiation role in informal proceedings between the teachers and the board of education regarding teacher welfare and/or personnel policy items, the study was concerned with the relationship of satisfaction to the superintendent's negotiation role as defined and restricted by a written contractual agreement between the two parties.

In order to examine these relationships within a formal negotiation atmosphere in contrast to informal teacher–board discussions, the following constraints were imposed. Only those districts reporting to the AASA (American Association of School Administrators), AFT (American Federation of Teachers), and NEA (National Education Association) a negotiated written contract including recognition, negotiation procedures, and provisions for impasse were considered for the study. A total of 272 school districts, representing twenty-six states were then selected according to the criteria stated. These districts ranged in size from 600 to 250,000 pupils.

The superintendent's negotiation roles ranged from one of full authority to negotiate for the board, to one of non-participation and are listed in Table 1 with the number of cases reported in each role.

As predicted prior to data collection, the middleman, (role 3) or advisor to both parties was the most frequent role assigned. This middleman assignment is due in large part to the historic positions of the NEA (4) and AASA (1). The dearth of superintendents in roles 1 and 2 may reflect current teacher opinion that they can themselves be effective spokesmen and do not want or need superintendents as intermediaries to the board. The infrequency of occurrence of role 6 along with findings reported later in this article, indicate that board members, in many cases, do not want to delegate much authority to superintendents but would rather control the power to act, using the superintendent in an advisory capacity.

Table 1.—Negotiation Role of the Superintendent

Role 1	Non-participant in negotiations	1
Role 2	Advisor to the Teachers' Organization only	0
Role 3	Advisor to both the Teacher's Organization and The Board of Education	109
Role 4	Advisor to the Board only	48
Role 5	Negotiator for the Board with limited authority	79
Role 6	Negotiator for the Board with full authority	14

The data for the study were collected from five participants to negotiations within each of the selected districts.

In addition to the superintendent, two board member negotiators and two teacher representative negotiators were randomly selected to participate from each district.

All respondents were asked informational questions regarding negotiations within the district including a forced-choice question on the negotiation role of the superintendent and the basis for this assignment, i.e., contract provision, board policy, board minutes, etc. The major portion of the questionnaire consisted of a number of statements developed according to Guttman's (6) recommendations for scale construction and presented to the respondents on the basis of a 5-point forced-choice scale reflecting a continuum of feelings from one of total satisfaction to one of total dissatisfaction. In this manner each respondent could be graded on a unidimensional scale of satisfaction as it pertained to the superintendent's negotiation role performance.

In the collection and analysis of the data, it quickly became evident that differences in role perception existed on the part of superintendents, teachers, and board members. The tabulation in Table 2 shows the extent of this conflict.

To assign the negotiation role in each district, the consensus of the three parties was used in categories one, two, and three. In those cases falling in categories four and five, the role assignment was judged, according to priority order, to be the role agreed on by (1) the majority of respondents, (2) a plurality of the respondents including the superintendent, or (3) the role agreed on by the board members.

The occurrence of such widespread conflict in role perception is of utmost importance to the three parties to the school negotiation scene. In order to lend stability to the bargaining situation as well as enhance the chance of positive outcomes, all parties ought to agree, regardless of the role chosen, on the role definition. Acceptance and understanding of the superintendent's negotiation role will allow a more complete use of the superintendent in negotiations and less misunderstanding of actions taken.

The data regarding the variables of school district size and the negotiation role of the superintendent were then subjected to the analysis of variance technique and a correlation coefficient was obtained. Because of the inconsequential number of cases in roles one, two, and six, only the data pertaining to roles three through five were used in this tabulation.

Table 2.—Superintendent Negotiation Role Perception Conflict

I	All parties in agreement	22%
II	Superintendent and Board Members in agreement	23%
III	Superintendent, one Board Member, one Teacher in agreement	6%
IV	Superintendent and Teachers in agreement	21%
V	All parties in disagreement with the Superintendent	28%

As can be seen in Table 3, these data proved to be significant beyond the .05 level with respect to both the f ratio (5) and Fisher's (5) transformation of the coefficient of correlation. This upholds the hypothesis that there is a significant relationship between the superintendent's negotiation role and the size of the school district. The relationship between the variables is such that superintendents in smaller school districts are perceived to have less authority in negotiations and in the larger school districts they are perceived to have greater authority.

The data collected from superintendent questionnaires were subjected to Guttman Scaling (2) in order to arrive at a satisfaction rating for each superintendent. In addition, to eliminate any possible distortion because of role perception conflict, the data were analyzed twice. In the first analysis all cases were used, and in the second analysis only those cases were used in which there was district consensus with regard to the superintendent's role. As can be seen from Table 4, neither analysis indicated the existence of a significant relationship between the variables.

Further analysis of the data, however, did reveal an increasing mean satisfaction for superintendents as the negotiation authority rose. Although a trend is apparent, the absence of a significant relationship may be partially explained by the amount of role perception conflict present and perhaps also by other factors affecting the superintendent's satisfaction.

The data collected from teacher representative respondents showed a significant relationship between the satisfaction of teachers and the superintendent's negotiation role. Using all cases, the analysis in Table 5 shows an f ratio beyond the .01 level of significance. The table further shows a comparison of the means via Duncan's Multiple Range Test (7). Means that are not underlined by the same line are to be considered significantly different.

The data show that teachers are significantly more satisfied when the superintendent is a nonparticipant to negotiations, less satisfied when he is in an advisory role to the board, and dissatisfied when he has authority for negotiations. These findings suggest the proposition that teachers

Table 3.—Superintendent's Role and Size of School District

	S.S.	df	M.S.	f
[a] Between Variance	84,927.538	2	42,463.769	3.049*
Within Variance	3,245,140.039	233	13,927.639	
Total Variance	3,330,067.577	235		

[a] Coefficient of Correlation r = .16*

* Significant beyond the .05 level

Table 4.—Superintendent Satisfaction with Negotiation Role

	S.S.	df	M.S.	f
A. All Cases				
Between Variance	55.964	4	13.991	1.368
Within Variance	2,516.768	246	10.231	
Total Variance	2,572.733	250		
B. Agreement Cases				
Between Variance	9.573	3	3.191	0.371
Within Variance	558.717	65	8.595	
Total Variance	568.290	68		

Table 5.—Teacher Satisfaction with Superintendent's Role

	S.S.	df	M.S.	f
[a]Between Variance	168.669	4	42.167	5.009**
Within Variance	2,744.183	326	8.418	
Total Variance	2,912.852	330		

[b]Role	Role 1	Role 4	Role 3	Role 6	Role 5
Mean	2.000	9.220	9.787	9.941	10.769

** Significant beyond the .01 level

want to deal directly with the power source and do not want to have the superintendent take an authoritative position in the structure of negotiations.

The data collected from board member negotiators were also analyzed. However, board members reported no cases in roles one and two and extremely few in role six (ten cases). For this reason, only three roles were used in the analysis; role three, role four, and a combined role five-six. The analysis of variance technique and the correlation of coefficient both indicated a relationship beyond the .05 level of significance (see Table 6). It also shows that role five-six is significantly different from role four, and role three is significantly different from role four.

These data reveal board member satisfaction with an advisory negotiation role for the superintendent. Board members are significantly less satisfied with both the middleman and the negotiator for the board approach. The research suggests that boards do not support the AASA position but prefer a superintendent who sits on the board side of the table. The data also suggest, however, that the board, while considering the superintendent's advice, still wishes to retain negotiation authority and deal directly with the new teacher power.

Summary and Implications

Keeping firmly in mind that the data analyzed reflect the formal negotiation role of the superintendent in contract districts, and that these data have been collected from predominately medium-sized school districts, let us attempt to generalize for the future benefit of those who will be forced to assume one such role.

A most important and most evident conclusion is that considerable conflict exists in the perception of the superintendent's negotiation role. Some evidence of this role perception conflict existed in 78 percent of the districts studied, and the conflict was serious in over 40 percent of the districts. It may be found that some superintendents can not assume certain negotiation roles because of prior background, training, and experience and hence create role ambiguity for themselves as well as encourage misevaluation of this role by others.

Table 6.—Board Members Satisfaction with Superintendent Role

	S.S.	df	M.S.	f
[a]Between Variance	19.049	2	9.525	3.304*
Within Variance	789.976	274	2.883	
Total Variance	809.025	276		

[b]Coefficient of Correlation r = .178

[c]Role	Role 4	Role 3	Role 5-6
Mean	8.327	8.874	9.049

*Significant beyond the .05 level

Regardless of the role chosen, however, the fact that a role definition is accepted and understood by the parties to negotiation will enhance the chances of positive outcomes for the superintendent as well as provide a measure of stability in the bargaining situation.

A second conclusion drawn from this research would be that there is a small but significant relationship between the assigned negotiation role and the size of the school district. Superintendents in large school districts have more authority in negotiations than superintendents in small districts. This is consistent with Bowman's (3) findings that superintendents in larger school districts tend to be more qualified, assume more responsibilities and prefer a more directive role than superintendents in smaller districts.

Although the data revealed an increasing mean satisfaction for superintendents as the negotiation authority rose, no significant relationship was indicated between the variables of negotiation role and superintendent satisfaction with the role.

In the case of teachers and board members, however, the study revealed a definite relationship between satisfaction and role. The teachers were most satisfied with the non-participant role, less satisfied with the advisory role, and dissatisfied with the negotiator role. The board members displayed a somewhat similar sentiment by being more satisfied with the advisory role than either the negotiator or the middleman role.

This dissatisfaction of these parties with the authoritative role may reflect the desire on both sides to do business on a first-line or toe-to-toe basis. Teachers may feel that they will do better if they bargain directly with the policy-makers, and board members may feel that a confrontation will persuade or reduce the opposition. Attitudes of this nature, if held, will produce inconsistent results, ranging from a peaceful cold war atmosphere to open strife, depending on the personalities involved, the negotiation skill available, and political pressures.

It is contended that such confrontations will soon lead boards to become dissatisfied because of the pressures, the extensive time required, and the administrative skill needed. Boards finding themselves in circumstances such as these will be likely to hire superintendents who are strong executives, and can resolve some of the personnel and negotiation problems without involving the board at every turn.

The stature of the superintendent of schools has already been affected by the new demands of a negotiation role, and only by further study of the variables and cogent application of the framework thus developed can the superintendent retain his position of educational leadership.

REFERENCES

1. American Association of School Admininstrators, *School Administrators View Professional Negotiations*, American Association of School Administrators, Washington, D. C., 1966, p. 55.
2. BMD 055, *Bio-Medical Computer Programs*, School of Medicine, University of California at Los Angeles, January, 1964, p. 379.
3. Bowman, Thomas, "Participation of Superintendents in School Board Decision-Making," *Administrator's Notebook*, January, 1963, p. 2.
4. Department of Classroom Teachers, *Classroom Teachers Speak on Professional Negotiations*, National Education Association 1967, Washington, D. C., p. 1.
5. Edwards, Allen, *Experimental Design in Psychological Research*, Holt, Rinehart and Winston, Inc., New York, 1960, p. 119.
6. Guttman, Louis, "A Basis for Scaling Qualitative Data," *American Sociological Review*, 1944, pp. 139–150.
7. Kramer, Clyde, "Extension of Multiple Range Tests to Gross Means with Unequal Numbers of Replications," *Biometrics*, 12:307–310, 1956.
8. "Negotiation Roles of the Superintendent," *1967 Progress Report of Contracts and Negotiations Under PA 298*, Connecticut Education Association, Hartford, March, 1967, Mimeo.
9. "State Patterns in Negotiations," *NEA Research Bulletin*, Washington, D. C., March, 1968.

New - - - - **ON TO THE CLASSROOM**

**A Guide for Preschool Teachers in
Head Start-Type Programs of Compensatory Education**

Edited by ROBERT E. CLASEN
200 pages $7.25 Hardcover and $5.75 Softcover

ORDER FROM—Dembar Educational Research Services, Inc.
Box 1148 • Dept. W • Madison, Wis. 53701

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

An Instrument for Evaluating Survey Research

RICHARD L. KOHR and MARILYN N. SUYDAM
The Pennsylvania State University

ABSTRACT

An Instrument for Evaluating Survey Research Reports was developed after synthesis of the suggestions of experts. A reliability study was conducted with articles from the field of elementary school mathematics education and raters from that field as well as from educational psychology and members of a research organization. Estimates of inter-rater agreement ranged from .80 to .95, while estimates of reliability for a single rater ranged from .34 to .86 for these judges. The instrument is suggested for use not only in evaluation but also in planning and developing research studies.

A CONSISTENT finding in all areas of research is the increased amount of published material, popularly known as an information explosion. According to Licklider (5) the amount of recorded scientific and technical information is increasing in an exponential fashion with the doubling rate ranging from 10 to 15 years. In addition, he noted that many low quality articles continue to appear in print.

It would seem to be of considerable importance for research workers, students, teachers, and administrators in education to become more critically oriented in their appraisal of published research. If research findings are to be used in determining educational practice and further research, some attempt to analyze and evaluate the various aspects of a report are vital.

Any systematized attempt to judge the 'goodness' of a research product should lead to an increase in the reliability of such judgments. Often the critical appraisal of a report, particularly by those who are inexperienced in this activity, is inconsistent in the sense that certain aspects are sometimes considered and sometimes not. For example, the reader may give attention to the sampling procedures for one study and fail to do so when evaluating another study. Systematized approaches to evaluation in the form of checklists or a similar type of instrument can serve to remind the evaluator of those aspects of research that he should be attending to consistently.

An instrument designed to evaluate experimental research reports was reported by Suydam (11). She gathered lists of suggestions for evaluating educational research from eighteen sources. In addition, she cited six reports which present instruments for such evaluation, only three of which included reliability data. From these twenty-four sources Suydam abstracted nine major aspects which were frequently and consistently repeated. Experimental research reports were evaluated by responding to each of nine resulting questions in terms of various "key points" which were subsumed under each question. Two reliability studies were reported in which coefficients of inter-rater agreement of .94 and .91 were obtained. Reliability estimates for a single judge were .77 and .57. Comparable values had been obtained for two of the other instruments cited by Suydam: Gephart (3), who found interrater agreement to be .74, and Johnson (4), who cited inter-correlations ranging from .75 to .79. From these data it would appear that the reliable use of instruments is possible.

Suydam's instrument dealt with experimental studies, and, like previous instruments, did not seem entirely appropriate for all other types of research. In particular, there is a need for an instrument for assessing survey reports. Several recent educational research texts (e.g., 12, 8) have indicated that surveys of various types are becoming more and more widely used for obtaining information not readily available by direct means. They also point out that surveys are increasingly used for administrative decisions. Often, however, such surveys are poorly designed and employ inadequate sampling procedures. In view of the generally agreed upon goal of improving research in education, it would seem that an important contribution would be the development of an instrument for evaluating survey research.

The instrument reported here was developed in connection with a research dissemination project dealing with elementary school mathematics.[1] Since this project is an extension of one previously reported (10), for continuity the definition of a survey used here follows that used in the original study:

> Research which attempts to find characteristics of a population by asking a sample through the use of a questionnaire or interview: includes also the status study, in which a group is investigated as it is to ascertain pertinent characteristics.

Survey research involves a body of procedures which may be used for the acquisition of information, primarily to determine the present status of a specified group. This characteristic distinguishes surveys from experimental studies in which the isolation of causative factors is the purpose. Surveys may be distinguished from historical research in that survey research deals mainly with present status as opposed to the past. Further, a distinction can be drawn between surveys and case studies. The former are usually conducted on cross-sectional samples while case studies typically involve the intensive and often longitudinal collection of information, with relatively small samples; they frequently seek to establish causal relationships.

The survey evaluation instrument was constructed by abstracting the major questions and the subsumed key points from numerous articles and books dealing with survey methodology. Presented below is the instrument for evaluating survey research reports, preceded by the directions for the instrument's use.

INSTRUMENT FOR EVALUATING SURVEY RESEARCH REPORTS

Richard L. Kohr
The Pennsylvania State University

Directions:

The following instrument is to be used for evaluating survey research reports within the framework of curriculum research. It is composed of nine major questions which are underlined. You are to rate the quality of the report in terms of *each* of these nine questions using the following 5-point scale:

1) Excellent: all requirements for the question are met; nothing essential could be added.
2) Very Good: most requirements are met.
3) Good: some requirements are met.
4) Fair: a few requirements are met.
5) Poor: none or too few of the requirements are met.

In determining a rating for each question, certain "key points" should be considered. These are listed below the question, followed by adjectives which indicate the continuum on which the key point should be assessed. Do NOT make a response to these "key points." They are intended to focus the attention of all raters on the same pertinent aspects of each question. In some studies certain "key points" may be irrelevant. In such cases base your judgment on those "key points" which are relevant. It is also possible that you may think of "key points" not included among those listed under a major question. Where relevant such additional "key points" may be used in assessing that question. There may be some instances in which none of the "key points" seem relevant or where the report fails to supply sufficient information. If this occurs, evaluate the question in terms of what you think should have been done and/or what information should have been included.

Please make only nine responses for each article, one for each question.

1. *How practically or theoretically significant is the problem?*	(1–2–3–4–5)
a. Purpose	(important—non-important)
b. Problem origin	
(1) Rationale	(logical—illogical)
(2) Previous research	(related—unrelated)
c. Generalizability	(extensive—limited)
2. *How clearly defined is the survey problem?*	(1–2–3–4–5)
a. Objectives and procedures	(specified—unspecified) (operational—vague)
b. Delimitations	(noted—not noted)
c. Variables	
(1) Control	(relevant—irrelevant) (operational—vague)
(2) Dependent	(relevant—irrelevant) (operational—vague)
3. *How relevant and how well defined is the population?*	(1–2–3–4–5)
a. Precise definition of population	
(1) Geographical limits	(specified—unspecified)
(2) Time period covered	(specified—unspecified)
(3) Sociological description	(specified—unspecified)
(4) Sampling units	(specified—unspecified)
b. Relevance of defined population to problem	(relevant—irrelevant)
4. *How adequate are the sampling procedures?*	(1–2–3–4–5)
a. Adequacy of sampling frame	

(1) Time period covered	(current—outdated)
(2) Inclusiveness of defined population	(complete—incomplete)
b. Method of sampling	(specified—unspecified)
c. Obtained sample	(appropriate—inappropriate)
(1) Size	(sufficient—insufficient)
(2) Representativeness	(adequate—inadequate)
5. *How adequately are sources of error controlled?*	(1–2–3–4–5)
a. Sampling error	(controlled—uncontrolled)
b. Non-response	(controlled—uncontrolled)
c. Interviewer bias	(controlled—uncontrolled)
d. Response error	(controlled—uncontrolled)
e. Response set	(controlled—uncontrolled)
f. Experimenter bias	(controlled—uncontrolled)
g. Teacher effect	(controlled—uncontrolled)
h. Control variables	(controlled—uncontrolled)
i. Extraneous factors	(controlled—uncontrolled)
j. Qualifications of research personnel (interviewers, coders, observers)	(experienced—inexperienced) (trained—untrained)
6. *How adequate are the measuring instruments?*	(1–2–3–4–5)
a. Choice of measurement technique(s)	(appropriate—inappropriate)
b. Instrument(s)	
(1) Development of instrument	(pretested—not pretested) (satisfactory—unsatisfactory)
(2) Description of administration and scoring procedures	(clear—unclear) (complete—incomplete)
(3) Wording of statements or questions	(clear—unclear)
(4) Sequence of statements or questions	(logical—illogical) (random—fixed)
(5) Evidence of reliability	(appropriate—inappropriate) (satisfactory—unsatisfactory)
(6) Evidence of validity	(appropriate—inappropriate) (satisfactory—unsatisfactory)
c. Rules for categorizing	(specified—unspecified)
7. *How appropriate is the statistical analysis of the data?*	(1–2–3–4–5)
a. Procedures of data collection	(specified—unspecified) (careful—careless)
b. Relation of obtained data to objectives	(essential—unessential) (sufficient—insufficient)
c. Descriptive measures	
(1) Statistic(s)	(appropriate—inappropriate)
(2) Evaluation of descriptive data	(appropriate—inappropriate)
(3) Establishment of relationships	(appropriate—inappropriate)
d. Statistical tests	
(1) Basic assumptions	(satisfied—unsatisfied)
(2) Relation to procedures	(appropriate—inappropriate)
(3) Significance levels	(specified—unspecified)
e. Description of results	(accurate—inaccurate)
8. *How reasonable are the conclusions drawn from the data?*	(1–2–3–4–5)
a. Interpretations	(consistent—inconsistent) (reasonable—exaggerated)
b. Generalizations	(reasonable—exaggerated)
c. Implications	(reasonable—exaggerated)
d. Qualifications	
(1) Discussion of methodological	

problems and errors	(comprehensive—limited)
(2) Alternative explanations	(noted—not noted)
(3) Other limitations	(noted—not noted)
9. *How adequately is the research reported?*	(1-2-3-4-5)
a. Organization	(excellent—poor)
b. Style	(clear—vague)
c. Grammar and mechanics	(excellent—poor)
d. Completeness	(excellent—poor) (replicable—unreplicable) (complete—incomplete)
e. Presentation of statistics	(clear—unclear)

Procedures

The population of survey research articles used in the present study were those in the field of elementary school mathematics (1900–1965) originally collected and reported by Suydam (10). The 230 articles were stratified according to year and journal. A stratified random sample of ten survey research reports were drawn. Xeroxed copies of these ten articles along with the survey evaluation instrument and notes to aid in its interpretation were given to nine judges who independently read and rated each report. Serving as judges were the following:

1. Three staff members of the Center for Cooperative Research with Schools (CReWS) who were experienced in the application of instruments in evaluating research reports and who were familiar with the content area of elementary school mathematics.
2. Three faculty members of the Educational Psychology Department. None of these judges were familiar with the content area nor did they have prior experience with instruments for evaluating research.
3. Three faculty members of the Elementary Education Department whose specialty area was elementary school mathematics. Each had prior experience in using instruments for the evaluation of research although not with the present instrument.

These nine judges may be regarded as fairly representative of staff members of a research organization who are experienced in research activities, and of research oriented faculty members of departments of educational psychology and elementary school mathematics.

Results

An internal consistency measure of reliability (inter-rater agreement) was determined by analysis of variance (ANOVA) procedures as outlined by Rabinowitz and Eikeland (9) and Medley and Mitzel (7). The ANOVA model used for determining reliability for each group of judges separately and for all judges combined is that of a randomized block design in which Articles is the block effect, Judges is the repeated measure, and the Articles × Judges interaction is the error term. Articles and Judges are both regarded as random effects in the present context. The dependent variable is the total score for an article derived by summing the score for the nine questions or items comprising the survey instrument. Reliability or inter-rater agreement is found by:

$$r_{tt} = \frac{MS_{\text{Articles}} - MS_{\text{Articles x Judges}}}{MS_{\text{Articles}}}$$

Since the intent is to discriminate among articles, the total variance of concern is that due to the Articles effect as shown in the above formula. True score variance is estimated by subtracting the error variance from this total variance. The resulting ratio of true score variance to total variance estimates the reliability for the situation in which all judges rate the articles. It is, of course, important to obtain a reliability estimate for a single judge or rater which reflects the common situation in which only one judge rates a study. This estimate of reliability for a single rater may be found by:

$$MS_{\text{True Score}} = \frac{MS_{\text{Articles}} - MS_{AJ}}{\text{Number of Judges}}$$

$$r_{tt} = \frac{MS_{\text{True scores}}}{MS_{\text{True scores}} + MS_{AJ}}$$

The value obtained by the above formula is the same as that obtained by Snedecor's formula described by Ebel (1).

Table 1 summarizes the reliability estimates obtained by separate analyses of variance for each of the three groups of judges and for all nine judges combined. A final reliability estimate is reported in which the effect of judges is separated into two effects—one due to groups of judges and the other as judges within groups. Thus, Articles × Judges interaction is partitioned into Articles × Groups and a residual (Articles × Judges within groups). This particular analysis is used for an additional purpose later in this investigation. In this design judges are nested within groups. This clustering of judges is termed strata by Rabinowitz and Eikeland (9), who indicate that when strata effects are present appropriate estimates of reliability may be obtained for both strata random and strata fixed assump-

Table 1.—Reliability Estimates for Separate Groups of Judges and for all Judges Combined

Group of Judges	N	Reliability ANOVA	Reliability Single Judge
Educational Psychology	3	.80	.57
Elementary Education	3	.84	.64
CReWS Staff	3	.95	.86
All Judges Combined	9	.80	.34
All Judges with Groups as Strata	9	.91	.78

tions. In the present context strata (groups of judges) is regarded as a fixed effect.

As may be noted, the CReWS group had the highest inter-rater agreement, followed by the elementary education and educational psychology faculty members. Of interest is the lower obtained reliability when all nine judges are pooled. Since the separate reliabilities are equal to or higher than that obtained for all nine judges combined, a strata effect is suggested. The final analysis considered strata (groups of judges) as a fixed effect and the increased reliability confirms the presence of a strata effect. This indicates that the judges within each group correlate more highly among themselves than with other judges. It should be recalled that all judges rated the articles independently and that the strata effect suggests similarity of criteria applied and not a group decision.

An additional analysis was performed in order to determine how "equivalently" each group of judges was using the Instrument for Evaluating Survey Research Reports. Medley (6) outlines four criteria which should be met in order to consider two tests (or measurement procedures) as equivalent (parallel forms). Under the assumption that true scores and error scores are normally distributed, two measurement procedures are regarded as equivalent if they have equal means, variances, standard errors of measurement and rank order subjects in the same way. An analogous condition holds in the present situation. Here equivalence refers to whether the various groups of judges are using the survey instrument in the same manner (as indicated by the various psychometric characteristics). Evidence for equivalence would be obtained in the present situation if the various groups of judges do not differ significantly with respect to mean scores on the articles, variances, and errors of measurement, and if the articles are rank ordered in the same way by each group of judges. It should be noted, however, that a failure to find significant differences is not sufficient to establish equivalence when the degrees of freedom are small. It may only reflect insufficient power to detect differences when the null hypothesis is false. Tests for homogeneity of variance are performed by forming an F-ratio based on the mean square for Articles, derived from the separate ANOVA's for each group of judges:

$$F = \frac{\text{Largest } MS_{\text{Articles}}}{\text{Smallest } MS_{\text{Articles}}}$$

Table 2.—Mean Squares Used for Tests of Homogeneity of Variance and Homogeneity of Measurement Error

Group of Judges	MS Articles	MS Articles X Judges
Educational Psychology	132.093	26.504
Elementary Education	159.870	25.459
CReWS Staff	107.985	5.674

Table 3.—Tests of Homogeneity of Variance Among Groups of Judges

Mean Square Ratio	F*	P
MS_{Articles} Ed. Psych. / MS_{Articles} Elem. Ed.	1.21	n.s.
MS_{Articles} Ed. Psych. MS_{Articles} CReWS Staff	1.22	n.s.
MS_{Articles} Elem. Ed. MS_{Articles} CReWS Staff	1.48	n.s.

*df = 9,9 for each case

In order to test for homogeneity of measurement error an F-ratio is formed by placing the largest MS_{error} in the numerator and the smallest MS_{error} in the denominator for each comparison. The error term is the MS for Articles × Judges interaction. Again, these mean squares are obtained in separate ANOVA's for each group of judges. Table 2 below summarizes the mean squares for each group of judges.

Table 3 summarizes the F-ratios for each test of homogeneity of variance. The degrees of freedom (df) for each test is (9, 9) since the df associated with the Articles effect in each ANOVA is 9. As may be observed none of the F-ratios were statistically significant; therefore, tentative evidence for homogeneity of variance was found.

In Table 4 the F-ratios for homogeneity of measurement error among groups of judges are presented. The df for each test is (18, 18) since the df associated with the Articles × Judges effect in each ANOVA is 18. Two of these comparisons are statistically significant. In each instance the CReWS group had significantly lower errors of measurement than the comparison groups.

The tests for homogeneity of group means and of function were obtained within the context of a single analysis of variance which is summarized in Table 5 below.

In this analysis the Groups main effect is nonsignificant, thereby supporting the requirement of homogeneous means. Homogeneity of function is assessed by testing for the presence of an Arti-

Table 4.—Tests of Homogeneity of Measurement Error Among Groups of Judges

Mean Square Ratio	F*	P
$MS_{\text{Articles x Judges}}$ Ed. Psych. / $MS_{\text{Articles x Judges}}$ Elem. Ed.	1.04	n.s.
$MS_{\text{Articles x Judges}}$ Ed. Psych. / $MS_{\text{Articles x Judges}}$ CReWS Staff	4.67	.01
$MS_{\text{Articles x Judges}}$ Elem. Ed. / $MS_{\text{Articles x Judges}}$ CReWS Staff	4.99	.01

*df = 18,18 for each case

Table 5.—Analysis of Variance Summary Table for Testing Equality of Group Means and Homogeneity of Function

Source of Variance	S.S.	df	MS	F	p
Articles	3075.73	9	341.75		
Groups of Judges	281.09	2	140.54	1.40	n.s.
Judges within Groups	603.20	6	100.53		
Articles X Groups	523.80	18	29.10	1.52	n.s.
Articles X Judges within Groups	1037.47	54	19.21		

cles × Groups interaction effect. Here again the resulting F-ratio was non-significant, providing evidence for homogeneity of function.

Evidence for equivalence seems to have been tentatively established for three measurement characteristics, namely, homogeneity of means, variance, and function. However, in the present situation only three judges comprised each group, and only ten articles were rated by each judge. With a small sample, there is reduced power for detecting differences, and it is possible that the obtained equivalence is partly a function of the small number of articles and judges. Homogeneity of measurement error was not obtained, and the groups cannot be considered equivalent in that regard.

Conclusions and Discussion

What have the analyses of the data on this test of the reliability of the Instrument for Evaluating Survey Research Reports shown? There are two substantive findings: (1) judges within each of three specified groups tended to be more consistent among themselves than with other judges, and (2) the groups of judges were using the instrument with a fair degree of equivalence. The reliability within groups is high enough so that it can be assumed that, especially with a period of familiarization training, its use in a controlled situation is plausible. For the CReWS staff, apparently the same criteria were being used; therefore, higher reliability resulted. The low measurement error for this group is also indicative of this increased agreement. With the other groups, even though they had the same general background (elementary education or educational psychology), varying or different criteria were apparently being applied; therefore, lower reliability was achieved.

This would seem to indicate that training is important. However, procedures used to test this instrument differed from those used to test the previous and similarly constructed instrument (13) in that an addendum to aid in interpretation was included. Whether or not the addendum aided the evaluator was unexplored and may be considered as a possible variable influencing reliability. However, the difference in reliability achieved by the two instruments can probably be accounted for by the differences in the type of articles being evaluated and by the different judges doing the evaluation. In addition, the number of judges in each case was small, thereby allowing for sampling fluctuation to play some role. Yet the questions remain: *Does the additional knowledge of factors confuse the evaluator rather than aid him? Is 'intuition' an important element in evaluation? How can we train judges? If the addendum is not sufficient, what is?* The effect of training has need of further exploration.

The two findings have implications for further use of the instrument. They suggest that groups of judges with certain characteristics in common, which might be considered appropriate to a given set of articles, could evaluate those articles in a rather similar fashion. Thus, if or when evaluation of a set of survey articles is desired, such judges could use the instrument independently, with a fair degree of confidence that their evaluations would not be in conflict with those which might be given by other judges with the same background.

Therefore, the Instrument for Evaluating Survey Research Reports can very feasibly be used in the many situations arising today where it is or could be of value to evaluate studies—for publication, for comparing usefulness of conflicting results, for abstracts, for research classes. The statistical tests of reliability are necessary but not sufficient to determine actual usefulness of the instrument. Of far more basic importance is the purpose for developing such an instrument. If we are ever to become more precise in developing and using research, such as instrument is necessary. It can serve the distinctly vital function of making everyone who comes into contact with a survey more aware of the essential factors to be considered in such research. The questions and key points help to identify the aspects of survey research which are most important. Thus it can serve as a guide: to the researcher in designing his study, and in writing a report of his work, and to the reader in evaluating the study.

The instrument has content validity: it encourages analysis or evaluation of what educational researchers for a half century have been saying are vital to consider. It is so structured that weighting of two components—sampling and measurement—is intrinsic. As we consider the increasing emphasis on the need to make educational research a more viable aspect in the process of education, what we can learn from studies done in the past should help to promote better methods in the future. What we can apply to future studies could help to bridge the gap between research and practice. The instrument is a valuable tool in evaluating research as one step in this process. Development and use of evaluation instruments are valuable for improving the use of research findings and for the development of valid research studies.

FOOTNOTE

1. "Interpretive Study of Research and Development in Elementary School Mathematics," Grant No. OEG-1-7-068592-0174, sponsored by the *Bureau of Research*, Research Utilization Branch, United States Office of Education.

REFERENCES

1. Ebel, R. L., "Estimation of the Reliability of Ratings," *Psychometrika*, 15:407–424, 1951.
2. Garvey, W. D.; Griffith, B. C., "Scientific Information Exchange in Psychology," *Science*, 156:1655–1659, 1964.
3. Gephart, W. J., *Development of an Instrument for Evaluating Educational Research Reports*, Cooperative Research Project No. S-104, U. S. Office of Education, 1964.
4. Johnson, G. B., "A Method for Evaluating Research Articles in Education," *Journal of Educational Research*, 51:149–151, 1957.
5. Licklider, J. C. R., "A Crux in Scientific and Technical Communications," *American Psychologist*, 21:1044–1051, 1966.
6. Medley, D. M., "A General Procedure for Testing the Equivalence of Two Tests," paper delivered at annual meeting of M.C.M.U.E., February, 1957.
7. Medley, D. M.; Mitzel, H. E., "Measuring Classroom Behavior by Systematic Observation," *Handbook of Research on Teaching*, Gage, N. L. (ed.), Rand McNally Co., Chicago, 1963, pp. 247–328.
8. Mouly, G. J., *The Science of Educational Research*, American Book Co., New York, 1963.
9. Rabinowitz, W.; Eikeland, M. H., "Establishing the Reliability of Tests with Clustered Items," *Pedagogisk Forskning*, pp. 85–106, 1964.
10. Suydam, M. N., "An Evaluation of Journal-Published Research Reports on Elementary School Mathematics, 1900–1965," unpublished doctoral dissertation, The Pennsylvania State University, 1967.
11. Suydam, M. N., "An Instrument for Evaluating Experimental Educational Research Reports," *Journal of Educational Research*, 61: 200–203, 1968.
12. Travers, R. M. W., *An Introduction to Educational Research*, Macmillan Co., New York, 1964.

BOOK REVIEWS—Continued from Page 58

volume a collection of the more important documents forming the basis for the history of American Education is served by this book. Like all such collections of readings, it has been necessary to confine some of the readings to a few pages. In a great many instances, however, the entire reference is included.

The author starts with the study of the more important aspects of education in England from which the theories and practices in the Colonial years of the United States were borrowed. Early policies of American education are discerned including: charity and early community backgrounds, "The Problem of Educational Development, Agarian Education, and Urban Problems"; the effect of religion and politics upon the neutralizing of American education; "The Common-School Revival" in the early part of the 19th century; and the "Styles of Learning in the 19th Century." In addition, attention is focused on the spread of education throughout the West. An excellent selection of readings is devoted to the various reform movements, particularly the Oswego Movement in elementary education. In section eleven there are excellent sources relative to the education of Blacks and the facilities for Blacks with reference to earlier facilities in the 19th century and the beginning of the 20th. Another section is devoted to new theories related to the elementary school subjects.

There are five selections on the immigrant family and the education of immigrant children, two sections on higher education, one on universities and foundations and then one other on "College Domination." The final three sections treat the increased enrollment in education and the problems it creates, the "Origins of the New Curricula" (a section altogether inadequate), and a completing section on "Race, Learning, and New Definitions of Community," which treats such topics as the Supreme Court decision, *de facto* segregation in the North, student activism, and neighborhood schools.

Harl R. Douglass
Professor, Consultant on Education
Boulder, Colorado

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 2, October 1970)

Taped Instruction Effects on Problem-Solving Skills of Seventh-Grade Children

FRANCIS SEKYRA III
Mississippi State University

M. RAY LOREE
University of Alabama

ABSTRACT

Thirty-six Ss randomly selected from an average intellectual ability group of seventh-grade students attending one Mississippi junior high school were exposed to six successive tape recorded problem-solving practice periods. Items contained in the tape recorded programs were designed to assess and implement training in problem solving. The number of correct responses constituted a product measure of problem solving. In addition to the problem-solving product measure, process measures were obtained on the problem-solving subskills classified as: (a) extraction of information contained in the statement of a problem, (b) retrieval of information, and (c) combining operations required to arrive at the solution of a problem. A 6 x 6 Latin square design with trend analysis was employed to ascertain whether practice in problem solving yielded significant improvement in all four problem-solving processes. The results indicated that increasing amounts of practice in problem solving, implemented through exposure to taped programs, yielded significant incremental S improvement on each of the four problem-solving variables investigated. The results agree with and support the findings of previous research.

PROBLEM SOLVING can be defined as the individual's unique efforts to reduce a knowledge gap between an initial state of affairs and a desired state of affairs. The types of behaviors involved in successful problem solving are of infinite variety, yet all of these approaches imply the existence of problems which are terminated successfully through thinking processes (5). Successful problem solving requires more than the mere possession or retrieval of knowledge. The problem solver must actively manipulate and structure information as he develops the model which yields the solution to the problem. The activity that occurs between the initial and desired states of affairs contains the key to the understanding of problem-solving behaviors (8).

Extensive theorizing regarding the nature of problem-solving processes has resulted in the development of several distinct theoretical orientations, each purporting to describe the processes involved in obtaining successful solutions to problems. Loree (9) has classified and outlined the major problem-solving conceptualizations as: *(a)* neobehaviorism, *(b)* cognitive structuralism, *(c)* computer models, *(d)* Piaget's developmental stages, *(e)* factorial approach, and *(f)* logical analysis of observational data.

Loree (9) examined the nature of problem solving and concluded that the product measure classified as the number of correct responses and the process measures classified as: (a) extraction of information contained in the statement of a problem, (b) retrieval of information, and (c) combining operations functioned as major variables involved in arriving at the successful solution of a problem. Conceptually these problem-solving variables can be related to the intervening processes suggested by other theories. Thus, internal mediating processes appear central to the development of a problem-solving model.

Loree (9) ascertained that students in grade 4 experienced verbalization difficulties. Anderson (1) shaped logical behavior in six- and seven-year-old human subjects, but he advised and urged that older subjects be used because first and second graders did not verbalize effectively. A study of group gains indicated a non-linear increment (9). Successful problem solvers demonstrated rapid gains up through grade 7. Subjects in grades 8 and 9 did very little better than subjects in grade 7. Thus, it appears that there is an optimal age for training in problem solving with the Loree technique.

Means and Loree (10) investigated the effects of asking germane questions directed towards development of behaviors characteristic of successful problem solvers on the improvement of problem-solving abilities. They conducted four individual training sessions with sixteen sixth-grade students of average intellectual ability. The subjects were asked to solve problems selected from standardized achievement tests for the appropriate age level under two experimental conditions. Each training session commenced with a "think aloud" period, followed by a second free response period with questions posed by the experimenter.

Means and Loree obtained significant increases for the extracting information and combining operations performances. Retrieval of information and overall improvement categories yielded insignificant data, but a positive trend seemed to be present. It was suggested that a longer training period might produce the desired significant results.

Time and cost impracticalities were major limitations to the actual classroom implementation of this technique. Means and Loree suggested that such training might prove feasible if presented via a programmed technique.

Method

Subjects. Ss, 36 Mississippi junior high school seventh-grade students, were randomly assigned to one of six experimental groups. All Ss attended the same school. Initially, ten students were assigned to each group to insure that statistical requirements were fulfilled when the investigation was concluded. When the experimental treatment terminated, each group was randomly reduced to six Ss for statistical analyses. Of the initial sixty students selected, eighteen failed to complete the training period for various reasons, but they were approximately equally distributed among the treatment groups.

The Ss were selected at random from a roll of all students in the seventh grade who functioned within an intellectual ability range of 85 to 115 IQ points as determined by the California Test of Mental Maturity (CTMM) or the SRA Primary Mental Abilities Test (PMA). The estimate of each child's level of intellectual functioning was obtained from permanent school records. Intellectual estimates of the Ss completing the investigation yielded a mean IQ of 103.9. The eighteen students who failed to complete the training period averaged 102.5 IQ points obtained on the CTMM or PMA.

Individuals within the average range of intellectual functioning tend to exhibit greater diversity in the quality of their problem-solving abilities than individuals who deviate in either direction from the mean (9). Seventh grade Ss appeared to function within the optimal range of problem-solving growth (9). Thus, the Ss had maximal potential to demonstrate improvement in their problem-solving skills.

Other than demonstrated intellectual ability and grade level, additional S restrictions were not required.

Instruments. It was necessary to develop a problem-solving training device based on items appropriate to the Ss' grade level. Standardized achievement tests were the sources of items which were used in this investigation. Items were obtained from the Sequential Test of Educational Progress, Science Forms 4A, 4B, 3A, and 3B (3); Metropolitan Achievement Test, Science Section Intermediate Battery (2); Stanford Achievement Test, Intermediate II Battery, Science Section (7); and the Otis–Lennon Mental Ability Test, Intermediate Level, Form JR (11). A total of thirty-six items was selected from the standardized instruments, and six items were randomly assigned to each of the six experimental training instruments.

Each standardized item was classified as the A segment of an instrument, and it served as the basis for the construction of three additional questions classified as the B, C, and D segments of the instruments.

The purpose of the A segment was to present a standardized problem to the Ss in an effort to promote basic problem-solving training. The data obtained from the A segment were assumed to be a problem-solving product measure classified as the number of correct responses. Emitting the correct responses required the S to use some portion of basic problem-solving ability as defined by Loree (9).

From the A segment, a related question was designed to provide training and investigate the Ss' ability to extract information contained within the statement of a problem, and the related question was classified as the B segment. Correctly answering the B segment required some element of the total process of understanding the problem posed and the ability to focus attention upon the key words or phrases in the statement of the original question. A second question, designated as the C portion, was constructed to promote training in the process retrieval of information. The C segment required the S to retrieve or recall information needed in problem solving.

A third question, based upon the original item, was constructed and classified as the D segment. The purpose of the D segment was to provide training in the problem-solving subskill combining operations. The D segment required the S to emit at least some of the behaviors classified under the problem-solving subskill combining operations. The combining operations processes involved are classifying, comparing, abstracting, analyzing, synthesizing, evaluating, and structuring information to solve a problem correctly.

Thirty-six standardized achievement items were each prepared with three appropriately related questions. The thirty-six standardized achievement items were pooled and randomly assigned to one of six instruments. Thus, six standardized achievement items and their three respectively related questions yielded a problem-solving instrument which consisted of a total of twenty-four questions. The six problem-solving training instruments were randomly identified as Question Series A, Question Series B, Question Series C, Question Series D, Question Series E, and Question Series F. The six instruments were prepared to provide training in the problem-solving product designated as the number of correct responses and the problem-solving processes classified as: (a) extraction of information contained in the statement of a problem, (b) retrieval of information, and (c) combining operations. The total of six different instruments was necessary to implement successive practice periods. Initial order assignment was random, and systematically rotating the sequence of instrument presentation controlled for inter-instrument differences.

Implementation of individual problem-solving training has been determined to be impractical for groups of Ss (10). To facilitate controlled group exposure, the six instruments were prepared in tape recorded programs with accompanying mimeographed instruction booklets for the S.

The 26 pages in the mimeographed booklets were 8½″ by 5½″ in size. The first page functioned to record the S's name, the date, the name of the experimenter (E), the question series, and the scores obtained. The second page provided a sample problem so that Ss would understand the nature of the task. The 24 subsequent pages contained one item of a question series. It was determined that one item per page would prevent Ss from referring to previous questions or reading ahead. The booklets were identical to the instrument tape scripts except that the instructions and answers were not printed. This provided an automated method whereby Ss were collectively exposed to training instruments which were designed to effect improvement in the Ss' problem-solving skills. Goldbeck and Briggs (6) regard automated teaching techniques as a method for bridging the gap between the problems of learning as studied in the laboratory and the problems of teaching in the classroom.

On any given tape, a S heard each item presented aurally via the tape recorder while he simultaneously read the problem in the mimeographed booklet. A 30 to 45 second interval was allowed for thinking. The S was then instructed to select the correct alternative from a list of four alternatives. Following the S's response, the correct answer and logical reasons for the solution were presented via the tape recorder. Thus, it was believed that Ss would experience training in problem solving and problem-solving subskills because of the nature of the instruments. All tapes were prepared by the same person so that vocal variability was reduced.

Procedure. Ss experienced six 1-hour training sessions according to the instructions on one of the six taped problem-solving programs. A 1-week period intervened between practice sessions. During the six experimental periods, the groups were located in separate rooms, and each group was exposed to one of the six taped problem-solving programs. Systematically varying the presented series of programs controlled for variation due to order of presentation. This control was necessary because the items and programs could not be proven to be equivalent. However, the observations for any given practice session were based upon all the items. Thus, any gains recorded could be considered to be a result of practice.

During the problem-solving practice sessions, the groups were under the guidance of six graduate students trained as Es. Each taped program was permanently assigned to one E, who worked with a different group on successive weeks in an effort to control E variables. Each E set up the tape recorder, distributed the mimeographed booklets, and issued instructions to Ss. Ss were told that this was a training project to see how well they could follow directions and solve problems. They were admonished to listen very carefully and respond immediately when told to do so. A practice problem was presented via the tape recorder, and E interrupted the tape to answer the Ss' procedural questions. The tape was then presented in its entirety. During the hour sessions, E circulated about the room to determine that Ss were responding appropriately, to prevent Ss from copying, and to determine that Ss were on the correct page. In all cases, E was not allowed to answer any questions for the duration of the experimental period. At the conclusion of the session, the record booklets were collected, and Ss were instructed to refrain from discussing the taped questions with their peers.

Results

Records of Ss were analyzed by counting the number of correct responses for each of the problem-solving subskills. Each correct answer received one point toward the S's score on each part of each program. Summaries of the raw scores are reported in Table 1.

A Latin square with Trend Analysis as suggested by Edwards (4) was employed to determine the level of significance yielded by the Ss' data on the number of correct responses product measure and the process measures: (a) extraction of information contained in the statement of a problem, (b) retrieval of information, and (c) combining operations. The Latin square design was employed to control for extraneous variation due to program difficulty, order of presentation, E influences, and external extraneous variables. Thus, any obtained variation should reflect only gains in problem-solving abilities yielded through practice via tape recorded training sessions.

Table 1.—Mean Scores for Six Practice Periods

Problem-Solving Variable	Periods 1	2	3	4	5	6
Number of Correct Responses	3.25	3.81	4.28	4.69	4.78	5.11
Extraction of Information	3.44	3.61	4.56	4.92	5.08	5.25
Retrieval of Information	3.25	3.00	4.05	3.92	4.25	5.06
Combining Operations	2.92	3.08	3.44	3.94	4.28	4.83

Two major research hypotheses regarding the problem-solving performance of Ss who participated in this investigation were evaluated. These research hypotheses, stated in null form, were:

1. Practice in problem solving implemented through exposure to taped programs did not yield significant improvement in the product measure classified as the number of correct responses emitted by the Ss.
2. Practice in problem solving implemented through exposure to taped programs did not produce significant improvement in the following problem-solving subskill process measures:
   a. Extraction of information contained in the statement of a problem.
   b. Retrieval of information.
   c. Combining operations.

The first null hypothesis stated that practice in

Table 2.—Summary of Analysis of Variance for Number of Correct Responses

Source of Variation	Sum of Squares	d.f.	Mean Square	F
Order	19.93	5	3.99	4.07*
Error (a)	29.33	30	.98	
Tapes	1.93	5	.39	.32
Periods	85.93	5	17.19	13.98**
Error (b)[1]	209.83	170	1.23	
Total	346.95	215		

[1] The sum of the squares for error (b) is the sum of the Latin square error sum of squares (27.83) and the pooled S's X periods sum of squares (182.00) with 20 and 150 d.f., respectively. The Latin square error mean is equal to 27.83/20 = 1.39. The S's X periods mean square is equal to 182.00/150 = 1.21. For the test of homogeneity of the two variances, F = 1.39/1.21 = 1.15 with 20 and 150 d.f., a nonsignificant value.

* $p < .05$.

** $p < .01$.

problem solving implemented through exposure to taped programs did not yield significant improvement in the number of correct responses emitted by the Ss. From Table 1, the means for the number of correct responses for the six practice periods were reported as 3.25, 3.81, 4.28, 4.69, 4.78, and 5.11. In order to test the first null hypothesis, Ss' scores were placed in a Latin square design. Six Latin squares were formed with rows corresponding to Ss, columns corresponding to successive practice periods, and cells corresponding to the number of correct responses obtained on questions drawn from standardized achievement instruments. Each Latin square represented an order of instrument presentation.

Complete analysis of the number of correct responses data is reported in summary form in Table 2. Inspection of Table 2 indicates that order of tape presentation and succesive practice periods yielded statistically significant effects. Successive practice periods yielded an $F\ (5, 170) = 13.98$, $p < .01$; and these data indicated an incremental function. Complete orthogonal trend analysis was performed yielding an $F\ (1, 170) = 70.21$, $p < .01$ to determine that a statistically significant linear component was present.

The hypothesis of no difference in the production of the number of correct responses during successive practice periods was rejected at the .01 level of significance. It was concluded that increasing amounts of practice through exposure to taped programs yielded significant improvement in the number of correct responses emitted by the Ss.

Null hypotheses 2a, 2b, and 2c were posed for the purpose of investigating the effects of practice with taped programs on the problem-solving subskills: (a) extraction of information contained within the statement of a problem, (b) retrieval of information, and (c) combining operations reported by Loree (9) and Means and Loree (10). Since there was an overall improvement in problem-solving ability, this improvement should be reflected by a similar improvement in the various subskills involved in problem solving.

Stated in null form hypothesis 2a predicted that practice in problem solving implemented through exposure to taped programs did not yield significant improvement in the problem-solving subskill extraction of information contained in the statement of a problem. The summary data obtained for the extraction of information variable are recorded in Table 1. From Table 1 the means for extraction of information contained in the statement of a problem for the six successive practice periods were respectively reported as 3.44, 3.61, 4.56, 4.92, 5.08, and 5.25. In order to test null hypothesis 2a, the Ss' scores on extraction of information were placed in a Latin square design. Six Latin squares were formed with rows corresponding to Ss, columns corresponding to successive practice periods, and cells corresponding to extraction of information scores. Each Latin square represented the order of instrument presentation.

Complete analysis of the extraction of information data is reported in summary form in Table

Table 3.—Summary of Analysis of Variance for Extraction of Information

Source of Variation	Sum of Squares	d.f.	Mean Square	F
Order	4.80	5	.96	1.12
Error (a)	25.65	30	.86	
Tapes	4.63	5	.93	.86
Periods	107.30	5	21.46	19.87*
Error (b)[1]	183.50	170	1.08	
Total	325.88	215		

[1] The sum of squares for error (b) is the sum of the Latin square error sum of squares (25.15) and the pooled S's X periods sum of squares (159.35) with 20 and 150 d.f., respectively. The Latin square error mean square is equal to 24.15/20 = 1.21. The S's X periods mean square is equal to 159.35/150 = 1.06. For the test of homogeneity of the two variances, F = 1.21/1.06 = 1.14, with 20 and 150 d.f., a nonsignificant value.

* $p < .01$.

Table 4.—Summary of Analysis of Variance for Retrieval of Information

Source of Variation	Sum of Squares	d.f.	Mean Square	F
Order	9.52	5	1.90	3.58*
Error (a)	15.93	30	.53	
Tapes	3.75	5	.75	.63
Periods	97.63	5	19.53	16.28**
Error (b)[1]	204.83	170	1.20	
Total	331.66	215		

[1] The sum of squares for error (b) is the sum of the Latin square error sum of squares (8.43) and the pooled S's X periods sum of squares (196.41) with 20 and 150 d.f., respectively. The Latin square error mean square is equal to 8.43/20 = .42. The S's X periods mean square is equal to 196.41/150 = 1.31. For the test of homogeneity of the two variances, F = .42/1.31 = .32, with 20 and 150 d.f., a nonsignificant value.

* $p < .05$.

** $p < .01$.

3. Inspection of Table 3 indicates that successive practice periods scores yielded an F (5, 170) = 19.87, $p < .01$. Complete orthogonal trend analysis of the extraction of information data yielded an F (1, 170) = 96.15, $p < .01$ to indicate that a statistically significant linear component was present.

The hypothesis of no difference in performance during successive practice periods on the extraction of information variable was rejected at the .01 level of significance. It was concluded that increasing amounts of practice implemented through exposure to taped programs yielded significant improvement in the problem-solving subskill extraction of information contained in the statement of a problem. Extracting information might consist of a variety of different processes, but the data obtained reflected growth in the Ss' ability to evoke some aspects of this problem-solving subskill.

Stated in null form, hypothesis 2b predicted that practice in problem solving implemented through exposure to taped programs did not yield significant improvement in the problem-solving subskill retrieval of information. The data obtained for the retrieval of information variable are summarized in Table 1. From Table 1, the means for retrieval of information scores over the six successive practice periods were respectively reported as 3.25, 3.00, 4.05, 3.92, 4.25, and 5.06. In order to test the third null hypothesis, the Ss' scores on retrieval of information were placed in a Latin square design. Six Latin squares were formed with rows corresponding to Ss, columns corresponding to successive practice period, and cells corresponding to retrieval of information scores. Each Latin square represented the order of instrument presentation.

Complete analysis of the retrieval of information data is reported in summary form in Table 4. Inspection of Table 4 indicates that successive practice period scores yielded an F (5, 170) = 16.28, $p < .01$. Complete orthogonal trend analysis of the retrieval of information data yielded an F (1, 170) = 75.04, $p < .01$ to indicate that a statistically significant linear component was present.

The hypothesis of no difference in performance during successive practice periods on the retrieval of information variable was rejected at the .01 level of significance. It was concluded that increasing amounts of practice implemented through exposure to taped programs yielded significant improvement in the problem-solving subskill retrieval of information. Retrieval or recall of information necessary to solve a problem might consist of a variety of different processes, but the data obtained reflected growth in the Ss' ability to evoke some aspects of this problem-solving subskill.

Stated in null form, hypothesis 2c predicted that practice in problem solving implemented through

Table 5.—Summary of Analysis of Variance for Combining Operations

Source of Variation	Sum of Squares	d.f.	Mean Square	F
Order	22.28	5	4.46	4.25*
Error (a)	31.44	30	1.05	
Tapes	13.44	5	2.69	1.21
Periods	98.00	5	19.60	12.73**
Error (b)[1]	261.33	170	1.54	
Total	426.49	215		

[1] The sum of squares for error (b) is the sum of the Latin square error sum of squares (31.44) and the pooled S's X periods sum of squares (229.89) with 20 and 150 d.f., respectively. The Latin square error mean square is equal to 31.44/20 = 1.57. The S's X periods mean square is equal to 229.89/150 = 1.53. For the test of homogeneity of the two variances, F = 1.57/1.53 = 1.03 with 20 and 150 d.f., a non-significant value.

* $p < .05$.

** $p < .01$.

exposure to taped programs did not yield significant improvement in the problem-solving subskill combining operations. The data obtained for the combining operations variable are summarized in Table 1. From Table 1, the means for the combining operations scores over the six successive practice periods were respectively reported as 2.92, 3.08, 3.44, 3.94, 4.28, and 4.83. In order to test the fourth null hypothesis, Ss' scores on combining operations were placed in a Latin square design. Six Latin squares were formed with rows corresponding to Ss, columns corresponding to successive practice periods, and cells corresponding to combining operations scores. Each Latin square represented the order of instrument presentation.

Complete analysis of the combining operations data is reported in summary form in Table 5. Inspection of Table 5 indicates that successive practice period scores yielded an F (5, 170) = 12.73, $p < .01$. Complete orthogonal trend analysis of the combining operations data yielded an F (1, 170) = 68.99, $p < .01$ to indicate that a statistically significant linear component was present.

The hypothesis of no difference in performance during successive practice periods on the combining operations variable was rejected at the .01 level of significance. It was concluded that increasing amounts of practice implemented through exposure to taped programs yielded significant improvement in the problem-solving subskill combining operations. The combining operations required to solve a problem might consist of a variety of different processes, but the data obtained reflected growth in the Ss' ability to evoke some aspects of this problem-solving subskill.

Tests of linearity of means of the data obtained during six taped sessions for each variable yielded nonsignificant F values, and it was assumed that the trends for the six instruments did not increase differentially.

Trend analysis of the extraction of information data yielded a significant quadratic component [F (1, 170) = 4.07, $p < .05$]. Graphic representation of the extraction of information data permitted the following visual impressions: (a) a sharp increment in performance occurred from period two to period three, and (b) a gradual improvement resulted during periods three through six without any apparent plateau effect. Separate curvilinear analyses of the data obtained from the remaining variables yielded nonsignificant values.

Discussion

The purpose of this investigation was to ascertain whether practice in problem solving effected through exposure to taped programs yielded significant improvement in Ss' problem-solving processes. The principal contribution of the present work was to ascertain the feasibility of the development of an automated, programmed method of training which would eliminate time and cost barriers which have heretofore prevented or limited classroom application of previous research efforts.

The results of the present experiment are in general agreement with the majority of previous studies assessing the effects of practice in problem solving. Examination of the data showed a definite incremental effect which was attributed to the successive exposure to taped programs designed to provide practice on four aspects of problem solving. The four constructs were the problem-solving product measure defined as the number of correct responses and the problem-solving process measures classified as: (a) extraction of information contained in the statement of a problem, (b) retrieval of information, and (c) combining operations as hypothesized by Loree (9). These data are in agreement with and support the previous findings of Loree (9) and Means and Loree (10).

The items contained in the taped programs probably did not effect significant knowledge growth. Rather, it was concluded that Ss developed greater proficiency in forming a problem-solving set. It was, for example, observed that Ss apparently solved the problems more rapidly and became bored during the pause interval allowed for problem solving. Thus, problem solving appears to be a skill which can be improved with proper practice conditions.

Extrapolation of these results should be limited to the grade placement and intellectual ability levels of the S involved in the investigation and the type of subject matter which served as a basis for the taped programs. It would be considered inappropriate to generalize these results to students in other grades or to different subject matter without additional research verification. Also, it should be noted that each problem-solving skill investigated consists of a variety of processes. It was not possible to ascertain which elements of the total process and in what degree these elements were involved in the problem-solving experiences generated by this investigation.

Statistically significant order effects obtained through analysis of the number of correct responses, retrieval of information, and combining operations data limit the generalizability of conclusions based upon this investigation of the particular Latin squares employed. The order of the first Latin square was generated according to random instrument selection. Subsequent orders were based upon systematic rotation of the initial order of presentation. Thus, it could not be concluded that the orders used in this investigation were representative of the entire population of possible six by six Latin squares. Therefore, it can be asserted that only the orders used in this investigation were controlled. The significant order effects cannot be attributed to unequal difficulty of the six tapes.

As a basis for further research, the present investigation should be expanded to include other grade placement levels, broader intellectual ability ranges, other subject matter, and questions other than objective achievement type problems. Length of the practice period should be manipulated to determine if there is an optimal duration for problem-solving training sessions. Also, full-sized classroom groups should be used to determine whether or not group size is a relevant variable.

Application of the technique developed during this investigation to the educative process could possibly provide a valuable tool to the classroom teacher. It should be noted, however, that no assumptions have been made regarding the long term effects of this type of problem-solving training. Permanency of problem-solving training should serve as a stimulus for additional research.

REFERENCES

1. Anderson, R. C., *Shaping Logical Behavior in Six- and Seven-year Olds, Cooperative Research Project No. 1790A*, University of Illinois Press, Urbana, Illinois, 1964.
2. Bixler, H. H.; Hildreth, G. H.; Lund, K. W.; Wrightstone, J. W., *Metropolitan Achievement Test, [1960 Edition]. Grades 5–6. Intermediate Battery, Form A*, Harcourt, Brace, and World, New York, 1958.
3. Cooperative Test Division, *Sequential Tests of Educational Progress: Science. Grades 4–6, 7–9; 1957; Forms A, B*, Educational Testing Service, Princeton, New Jersey, 1957.
4. Edwards, A. L., *Experimental Design in Psychological Research*, Holt, Rinehart, and Winston, New York, 1960.
5. Gagné, R. M., *The Conditions of Learning*, Holt Rinehart, and Winston, New York, 1965.
6. Goldbeck, R. A.; Briggs, L. J., "An Analysis of Response Mode and Feedback Factors in Automated Instruction," in Smith, W. I.; Moore, J. W.; (eds.), *Programmed Learning*, D. Van Nostrand, Princeton, New Jersey, 1962, pp. 166–191.
7. Kelly, T. L.; Madden, R.; Gardner, E. F.; Rudman, H. C., *Stanford Achievement Test, [1964 Revision]. Grades 5.5–6.9, 7–9*, Harcourt, Brace, and World, New York, 1964.
8. Loree, M. R., *Psychology of Education*, Ronald Press, New York, 1965a.
9. Loree, M. R., *Problem-solving Techniques of Children in Grades Four Through Nine. Cooperative research project No. 2608*, Research Report to the U. S. Office of Education, University of Alabama Press, University, Alabama, 1965b.
10. Means, G. H.; Loree, M. R., "The Influence on Problem-solving Ability of Following an Efficient Model, *The Journal of Educational Research*, 1968, 62:(3), 135–141.
11. Otis, A. S.; Lennon, R. T., *Otis–Lennon Mental Ability Test, Intermediate*, Grades 7–9; Harcourt, Brace, and World, New York, 1966.

FIELD NEWS

Willard E. North
Central Missouri State College

AN ERIC CLEARINGHOUSE SPECIALIZING in evaluation and measurement was established at Educational Testing Service (ETS) in May 1970. The new Center will be operated in association with Rutgers University Graduate School of Education, and will be under the direction of Henry S. Dyer. Funded by USOE, the Educational Resources Information Center for Tests, Measurement, and Evaluation was created to provide research psychologists and other educators with a central source of information on research in these areas, and at the same time helps to stimulate additional research.

Plans call for the Center to act as a depository of information on new evaluation and measurement materials and procedures, to process and analyze pertinent materials and techniques, and to make its findings available to educators through publications, on-site services, and, ultimately, regional conferences. The Clearinghouse staff will be aided by members of the Rutgers faculty and by ETS research groups and other resources of the organization. An Advisory Board, representing broad sections of the educational community will help provide guidance for the Center's development. The new Center joins nineteen other ERIC Clearinghouses that have been created by the USOE to disseminate results of new educational materials, procedures, and research.

* * * * * *

THE EXPERIENCE OF 129 4-YEAR MIDWESTERN colleges and universities in recruiting and enrolling minority students (Blacks, Indians, and Spanish–Americans) is reported in the first of several Higher Educational Surveys of current problems in college admissions planned by the College Entrance Examination Board. This survey report, titled "Admission of Minority Students in Midwestern Colleges," summarizes information collected in April 1970 from a representative sample of Midwestern senior institutions on such topics as enrollment and recruitment of minority students, retention rates, financial aid, and the outlook for the immediate future.

Designed to provide its member institutions with a rapid means of gathering information which is timely and relevant to current problems in higher education, the Higher Education Surveys are a new experimental project of the College Board. The "Admission of Minority Students in Midwestern Colleges" was executed by the Midwest Committee for Higher Education Surveys with staff support from the Board's Midwest Regional Office in Evanston, Illinois, and Access Research Office in Palo Alto, California.

Among the principal findings of this first survey are:

1. Enrollment of minority students in Midwestern senior institutions is increasing. Admission officers report a 25 percent increase from 1968 to 1969 and estimate an additional increase of 30 percent this fall. Three-fourths of the colleges anticipate further increases over the next 3 years.
2. Three-fourths of the public selective institutions judge their minority freshmen to be below the general freshman average in academic potential. A substantial majority of all institutions having minority freshmen of below average potential are trying to give them special academic help.
3. Approximately seven out of ten minority freshmen return to their institution the following year. This is the same proportion reported for all freshmen.
4. The proportion of minority students requiring full financial aid is more than three times as high as that for all incoming freshmen. Admission officers estimate that some 15 percent of all freshman financial aid is awarded to those 5 percent who are minority students.
5. New recruiting techniques which seem to be most successful include using minority staff and students, special programs for minority students, and direct contact with schools.

Copies of the full report may be obtained from: College Entrance Examination Board, 625 Colfax Street, Evanston, Illinois 60201.

* * * * * *

THE GROWING NUMBER OF COOPERATIVES being set up by the nation's academic libraries to meet increasing costs and demands for services is a subject of a study by the System Development Corporation (SDC). Working under a 16-month, $92,000 contract from the USOE, SDC is studying existing consortia (cooperative networks) and will produce a directory identifying all known academic library cooperatives, listing participating libraries and services.

This directory will provide the basis for the project's second phase: in-depth study of fifteen selected library cooperatives to help develop the guidelines that can be used by libraries to plan networks and other cooperative library arrangements. The guidelines will provide detailed, step-by-step information for planning, developing, operating, and evaluating academic library consortia. Dr. Carlos Cuadra is the project leader and head of SDC's Library and Documentation Systems Department.

* * * * * *

TEACHER'S COLLEGE PRESS HAS RECENTLY published "Educational Investment in an Urban Society." The book is coedited by Melvin R. Levin, Director of the Division of Urban Studies, and Alan Shank, Assistant Professor of Government, both at Boston University. This volume presents views of noted economists, political scientists, and public officials on the uses to which cost-benefit analysis has been put in education, and discusses the limitations and potential of the technique. Selections in this book cover the most recent developments on such topics as human capital, educational investment and productivity, cost-benefit analysis in education, planning-programming-budgeting systems, and social indicators. Among the contributors are Theodore Schultz, Gary Becker, Burton Weisbrod, Jerry Miner, Charles Benson, John Vaizey, Thomas Ribich, Jack Wiseman, Aaron Wildavsky, Warner Hirsch, Melvin Levin, Wilbur Cohn, and Star Levitan and Garth Mangum.

"Educational Investment in an Urban Society; Cost, Benefits, and Public Policy" is published by Teacher's College Press, Teacher's College, Columbia University, New York, New York 10027.

HE RESULTS OF A STUDY by the Bureau of College Evaluation on Changes in Academic Calendars are reported in the publication SHAPE issued by the Bureau of Research in Higher and Professional Education of the New York State Education Department. General findings of this survey are that changes in the academic calendar are increasingly common. In a survey of 197 New York State institutions of higher education, more than half have changed within the past 5 years or are currently considering change. The most common calendar configuration, however, is still some form of the semester system.

The popular trend in calendar modification is the interterm or "mini-semester" plan. This plan establishes a short session, usually about 4 weeks, in which a student works intensively upon one course or project. Time is made available for the "mini-semester" in January by ending the first semester before Christmas vacation and beginning the regular second semester in February. Proponents of "mini-semester" stress that its short intensive nature encourages curricular innovation, teaching experimentation, and student involvement.

* * * * * *

A DIRECTORY OF FEDERALLY SUPPORTED Information Analysis Centers is now available. Sponsored by the Panel on Information Analysis Centers (Panel 6) of the Committee on Scientific and Technical Information (COSATI), Federal Council for Science and Technology, the Directory provides useful information on 119 federally supported Information Analysis Centers, including name, address, and phone number; name of director(s) or operator(s), sponsor(s); year started; size and make-up of staff; mission and scope; services, and who may qualify as a user. Four indices are included as appendices: (A) Center Operators or Directors Index; (B) Geographic Index; (C) Organizations Index; and (D) Subject Index.

The new publication updates as of January 1970 a previous directory issued by COSATI in April 1968. Date of collection, indexing, and production of the revised publications were performed by the National Referral Center for Science and Technology of the Library of Congress, publishers of the familiar series *A Directory of Information Resources in the United States*. Some of the data for the COSATI Directory came from the comprehensive inventory of information resources maintained by the National Referral Center. The Directory may be obtained in either paperbound or microfiche form; the price is $3.00 or $.65 respectively, catalog number PB189300. Orders should be sent to Clearinghouse, U. S. Department of Commerce, Springfield, Virginia 22151.

* * * * * *

COPIES OF THE 300-PAGE REPORT, "The Arts, Education, and the Urban Sub-cultural," relating the findings and recommendations of the 2-year National Survey of the Arts for Minority Youth, are now available from the Communications Foundation, 2020 Alameda Padre Serra, Santa Barbara, California 93103 at $9.50 per copy. The report lists almost five hundred performing arts centers, including forty-six black arts projects, and indexes original films, scripts, and anthologies by student artists as resources for the arts educator.

* * * * * *

THE NATIONAL BOOK COMMITTEE HAS received a grant of $81,000 from the USOE to carry out Phase II of a nationwide Educational Media Selection Centers Program. The 4-phase program is designed to help develop media selection centers for use by classroom teachers, curriculum specialists, librarians, and others concerned with educating children and youth.

In Phase II a guide will be prepared for establishing or developing such centers. It will be based on an analytic survey of existing centers made during the initial research phase of the program. The findings of Phase I are to be published by the American Library Association this fall, under the title "Educational Media Selection Centers: Identification and Analysis of Current Practices."

The guide to be developed for the Educational Media Selection Centers will cover major aspects of their objectives, services, collections, and facilities. Included will be plans for the networks of centers; principles for evaluation, selection and acquistion of media; prototypes of advisory services and education programs for a variety of youth-serving adults; specific recommendations related to budget, space, administration, staff, and dissemination of information about availability of services of centers. The guide is to be developed during the next 12-month period. It is anticipated that publication of the guide will bridge a current information gap for agencies planning new centers or administrators seeking to upgrade and extend their existing center collections and services.

During Phase III, it is planned that model or demonstration centers will be established. They will be fully functional and will be located in different parts of the country; they will reflect a variety of administrative and service patterns. Phase IV is designed to evaluate both model centers and effectiveness of the Phase II guide. The Phase IV report will be augmented by a comprehensive information campaign and production of a film about programs and services.

Miss Cora Paul Bomar, Assistant Professor of Education, University of North Carolina at Greensboro, is director of the project.

* * * * * *

NONPROMOTION IN OUR SCHOOLS: A SIGN OF FAILURE, written by Athol B. Packer, Department of Elementary Education, University of Florida has been published as Research Bulletin (Volume 5, Number 1) by the Florida Educational Research and Development Council.

The bulletin explores studies showing the extent of failures in a number of school systems. It uses the findings of research to answer a number of questions which professionals raise about repeating a grade or subject. Among them are: (1) Does nonpromotion reduce achievement level? (2) Is academic achievement higher in schools with a high nonpromotion rate? (3) Does the individual child increase his academic achievement when he repeats his school grade? (4) What are the emotional and psychological effects of nonpromotion? (5) Is nonpromotion related to discipline problems? (6) Do pupils who are not promoted tend to drop out of school? (7) Is there a relationship between the rate of pupil retention and teacher professional activities, attitudes, and viewpoints? (8) What are the underlying causes of nonpromotion? The final section offers a number of suggestions for eliminating failures.

Copies may be obtained for $1.00 per copy with a 25 percent discount to libraries. Orders should be directed to: Dr. J. B. White, Executive Secretary, Florida Educational Research and Development Council, College of Education, University of Florida, Gainesville, Florida 32601.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

HAROLD N. BEYER is a doctoral student in the department of psychology at Colorado State University. He is presently counseling part time in the University Counseling Center while completing his graduate work. He received his BA from Sacramento State College and MS from Colorado State University. **EUGENE R. OETTING** is a professor of pychology at Colorado State University. He received his BS, MS, and PhD from The University of Wisconsin.

ERNESS B. BRODY is an assistant professor of education at the Graduate School of Education, Rutgers University. She was on leave in London, England, studying teacher training and English infant schools for the 1969–70 school year. During 1968–69 she was Director of the Experiment in Teacher Education at Rutgers. She received her BA from Fisk University and her MA and PhD from The University of Michigan. She is mainly interested in research on teaching and urban problems.

WILLIAM E. CALDWELL is associate professor of education at The Pennsylvania State University, where he is engaged in research and teaching in the field of teacher–school-board negotiations and personnel management in education. Former positions held have been mathematics teacher and coach, secondary counselor, school administrator, and assistant professor of secondary education. He earned the PhD at New York University after obtaining prior degrees at Eastern Nazarene College and University of New Hampshire. As Executive Secretary of The Pennsylvania School Study Council, he assists in the coordination and direction of research activities between the University and school districts. He is presently working on the construction of a computer simulation for use in training school administrators in personnel management.

RONALD P. CARVER received his PhD in experimental psychology from Washington University. He is currently Research Scientist with the American Institutes for Research in the Washington, D. C. office. His major research interest is in the processing and storage of meaningful verbal information. He is also interested in research design, educational measurement, test development and statistics. He has developed a new type of reading test called "chunked," and has investigated pupilary dilation as a measure of information processing load.

DOLORES DURKIN is a professor of elementary education at the University of Illinois, Urbana. She has been a consulting editor for *The Journal of Educational Research* since 1968.

MICHAEL J. A. HOWE is associate professor of educational psychology at the University of Alberta. He received his BA and PhD from the University of Sheffield, England, and was previously assistant professor of education at Tufts University. His publications include a number of articles and a book on human memory, as well as several articles concerning meaningful learning.

RICHARD L. KOHR holds an MS in educational research from Bucknell University and a PhD in educational psychology from The Pennsylvania State University. He was a research assistant with the Center for Cooperative Research with Schools and is currently a lecturer at Bucknell University. **MARILYN N. SUYDAM** is associate professor of education at The Pennsylvania State University and associate director of the Center for Cooperative Research with Schools. She received the BS and MS from State University College at Plattsburgh, New York, and the DEd from Penn State. She has been directing projects related to the compilation, evaluation, and dissemination of research on mathematics education.

BETTY RUTH RAYGOR is an assistant professor in education at Hamline University, St. Paul, Minnesota. She received her BEd from the University of Toledo, MA and Specialist in Education from the University of Minnesota. She was formerly Director of Title I and Coordinator of Special Education in the Roseville Public Schools, St. Paul.

FRANK SEKYRA functions in the capacity of associate professor and acting head of Mississippi State University's Department of Psychology. His BS was conferred by the University of Illinois, and MA and PhD were earned from the University of Alabama. Currently, Dr. Sekyra's research interests focus upon assessment of intellectual functioning and appraisal of intellectual assessment instruments. **M. RAY LOREE** is professor of educational psychology and director of the graduate research training program at the University of Alabama. His BA is from the University of Manitoba, and his MA and PhD from the University of Chicago. He is the author of *Psychology of Education.*

KARL SMITH is director of the Behavioral Cybernetics Laboratory, The University of Wisconsin. He obtained his AB from Miami University, Oxford, Ohio, and his MA and PhD from Brown University, Providence, Rhode Island. Dr. Smith previously taught at Brown University, and the University of Rochester. He was also Ford Foundation Distinguished Professor of Business Administration, Indiana Univerity. **HARVEY SUSSMAN** is a graduate student at The University of Wisconsin working on his PhD in the field of Behavioral Cybernetics. He earned his undergraduate degree at Queen's College, New York, and his master's at The University of Wisconsin.

Diary No 204
Date 9.2.71
File No S. K B.
[illegible] Ednl. Psy. Research

50th Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **NOVEMBER 1970** **NUMBER 3**

ARTICLES

A Decade of Research in Handwriting: Progress and Prospect—*Eunice Askov, Wayne Otto, and Warren Askov* -- 99

The EPPS Need Structure of In-Service Elementary School Teachers—*E. Larry Eberlein* -- 112

A Model for Teaching Composition—*Lester S. Golub* -- 115

Effects of Training on the Instructional Behaviors of Student Tutors—*Fred C. Niedermeyer* -- 119

Interaction Effects of Sex of Mother's Child and Child's Reading Performance on the Mother's Evaluation of the School—*Oren Glick* -- 124

Computer-Assisted Planning and Scheduling of Individualized Programs of Study in Science and Mathematics at the Secondary Level—*Sister Marijane Werner* -- 127

Teacher Personality and the New Social Studies—*James R. Robertson and John D. Haas* -- 133

Test Difficulty as a Factor in Achievement—*James R. Daines* 139

DEPARTMENTS

Among Those Present -- 144

Book Reviews—*Robert E. Clasen* -- 114

Field News—*Willard E. North* -- 142

Forthcoming Articles -- 123

It Says in The JER -- 98

Research Brief—*Frank B. Murray and Joseph R. Jenkins* -- 141

(The Journal of Educational Research is indexed in Abstr.S.W., CIJE, CSPA, Current Contents, Ed. Adm. Abst., Educ. Ind., M.H.R.Ind., PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1970 by Dembar Educational Research Services, Inc.
Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.
Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, Supervisor of Editorial Services.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor Curriculum and Instruction, The University of Wisconsin, Madison.

PHILIP LAMBERT, Chairman, Department of Educational Psychology and Administration, The School of Education, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1970

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Assistant Professor of Special Education, Yeshiva University, New York.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English and Education, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLASS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

CARLETON W. WASHBURNE, Director Emeritus of Teacher Education and the Division of Graduate Studies, Brooklyn College, 2248 Kent Street, Okemos, Michigan 48864.

It Says in The JER...

INSTRUCTIONAL PRACTICES in handwriting were found by Askov, Otto, and Askov to be based primarily on traditional procedures rather than on research findings. They encourage a reevaluation of the directions of handwriting research and of current instructional procedures with the intent that the 1970–80 decade may produce more definitive information in the study of handwriting (p. 99).

IN FEMALE elementary teachers, manifested needs change with age and experience, says Eberlein. The Edwards' Personal Preference Schedule was administered as a means of gathering data for this study (p. 112).

WHAT EFFECT DOES oral discourse have in teaching written discourse? Golub found that the use of controlled oral discourse produced more grade increases and fewer decreases in written discourse in the experimental group which went from stimulus to problem to controlled oral discussion to written discussion on the composition model. The control group went from stimulus to problem to written discussion with no oral discussion (p. 115).

FIFTH-GRADE TUTORS WERE used to help kindergarteners in a reading program. In Niedermeyer's study six of the tutors were trained and six untrained. Trained tutors scored significantly higher on engaging the pupil in friendly conversation, confirming, praising, giving the correct answer when the pupil is incorrect, and eliciting correct responses before going on (p. 119).

DO MOTHERS OF GIRLS react differently than mothers of boys in evaluating the school their child attends? Glick found that if the children fell below the reading criterion, the sex of the child made no difference in the mother's rating. For mothers of children performing above the criterion, mothers of girls evaluated the school much more favorably than mothers of boys (p. 124).

PLANNING AND SCHEDULING of math and science programs at the secondary level can now be assisted by a computer. Sister Marijane Werner found that knowledge of Program Evaluation and Review Technique (PERT) concepts and principles makes for more efficient preplanning of goals and objectives than might otherwise be the case. Used concurrently and consistently with other aspects of curriculum planning, PERT reveals interdependencies and problem areas which are neither obvious nor well defined by conventional planning and scheduling methods (p. 127).

PRESENT SOCIAL STUDIES teachers can deal with reflective inquiry in the classroom and handle the heat and pressure that characterize societal disputes which become the focus of classroom discussion, Robertson and Haas cautiously concluded. They measured the manifest needs of social studies teachers from Utah with the Edwards' Personal Preference Schedule (p. 133).

GAINS ASSOCIATED WITH the "low difficulty testing approach" were significantly higher than those of the "high difficulty testing approach." Daines investigated the effect of two testing approaches on informational achievement in a technical industrial education course (p. 139).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Editorial/Abstract

A Decade of Research in Handwriting: Progress and Prospect

◆ Early in the 1960's, Herrick and Okada identified various directions that research related to handwriting might take during that decade. Now the authors are attempting to consider both the extent to which research has actually been focused in these areas during those 10 years and the conclusions that can be drawn from the research.

Handwriting studies have been classified under the following categories which correspond to Herrick and Okada's research areas with a few slight modifications:

1. *The nature of letter forms most efficiently and legibly produced.*
2. *Instructional techniques most conducive to children's learning and perceptual-motor development, and correlates of skill development in handwriting.*
3. *Effects on writing performance of (a) specific positions of the body, feet, arm, wrist, fingers, and paper; (b) grip employed in holding the writing instrument; and (c) arm movement.*
4. *Effect of speed and stress on the handwriting product.*
5. *Effect of instructional sequences on the development of handwriting skills.*
6. *The nature of handwriting instruments and writing surfaces that facilitate learning to write.*
7. *Development of new handwriting scales and evaluative materials, and means for helping teachers to make better use of existing scales.*

In each area the findings from the various studies were compared. Conclusions and implications for handwriting instruction and research were drawn where justified. Often definitive conclusions could not be drawn, however, due to methodological differences in the studies.

The authors found that although research had been done in all of the above areas during the 1960–70 decade, the studies were concentrated in a few areas while other areas were relatively neglected. Instructional practices were also found to be based primarily on traditional procedures rather than on research findings. The intent of the article is to encourage a reevaluation of the directions of handwriting research and of current instructional practices with the intent that the 1970–80 decade may produce more definitive information in the study of handwriting.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

A Decade of Research in Handwriting: Progress and Prospect

EUNICE ASKOV, WAYNE OTTO, and WARREN ASKOV
The University of Wisconsin

SURVEYS of the research and literature related to handwriting have been reported by Andersen (1, 2), Herrick (35), Horn (43), Otto and Andersen (59), Petty (65), and Shane and Mulry (74). Since, as a group, they are both recent and thorough, there is no attempt here to replicate these existing surveys. The purpose, instead, is to examine the 1960–70 decade of research in handwriting to seek implications for instruction and directions for future research.

Early in the 1960's, Herrick and Okada (39) identified various directions that research in handwriting might profitably take during the decade. In 1970, it seems appropriate to consider both the extent to which the research has been focused in these areas and the conclusions that seem justified. This paper, then, is organized around the research areas suggested by Herrick and Okada. Some minor modifications were made in order to enhance the "goodness of fit" between the categories and the studies actually reported. Thus, there is an attempt to focus upon:

1. *The nature of letter forms most efficiently and legibly produced.*
2. *Instructional techniques most conducive to children's learning and perceptual-motor development, and correlates of skill development in handwriting.*
3. *Effects on writing performance of (a) specific positions of the body, feet, arm, wrist, fingers, and paper; (b) grip employed in holding the writing instrument; and (c) arm movement.*
4. *Effect of speed and stress on the handwriting product.*
5. *Effect of instructional sequences on the development of handwriting skills.*
6. *The nature of handwriting instruments and writing surfaces that facilitate learning to write.*
7. *Development of new handwriting scales and evaluative materials, and means for helping teachers to make better use of existing scales.*

The Nature of Letter Forms Most Efficiently and Legibly Produced

Research related to letter forms has taken several directions since 1960. One approach that is in line with the early, classic studies by Pressey (67) and Newland (57) was to study illegibilities to determine which letter forms are the most difficult to produce correctly. Lewis, in his dissertation study (48, 49, 50), identified not only the most frequent types of errors in first graders' manuscript writing but also the letters that were the most susceptible to error. The results provide specific guidelines for primary classroom teachers' efforts in the area of manuscript writing.

Approaching the problem from a different perspective, Schell and Burns (72) studied the modifications in uppercase cursive letters that were made by college students who had been exposed to the same instructional program in handwriting in the elementary school. Deviations tended in the direction of simplification of the letter forms that had been taught. The conclusion was that simplified letter forms ought to be taught initially to preclude the need for subsequent, and perhaps too highly idiosyncratic, modifications. The Schell and Burns results are, however, put in perspective by the findings of another study. Epstein, Hartford, and Tumarkin (18) found that female adults who retained the letter forms as they learned them in elementary school tended to be different from female adults who modified letter forms in that the former were not as well educated, bright, or mature. Since Schell and Burns studied college seniors, their sample may have represented mainly that portion of the population which would tend to modify the letter forms they were taught in elementary school.

In another related study, Seifert (73) examined the development of personal styles of handwriting by 1,020 children in grades 6–9. Only one-third of the children at each grade level gave evidence of having developed personal styles of handwriting, but the author concluded, on the basis of her analyses of writing samples, that their handwriting was legible though different from the standard taught. Interestingly, in most Ss no significant relationship was found between intelligence and the tendency to develop a personal handwriting style.

The main conclusion to be drawn from this group of studies is that the letter forms used in instructional programs might profitably be simplified without loss of legibility. (An early study by Quant (68) showed that simplified letter forms were in fact more legible.) Modifications of the letter form models employed in initial instruction could make for more efficient instruction and learning and result in more legible adult writing. Much of the basic data required in order to make such modifications has already been reported.

Some attempts have been made to use modified letter forms in instructional programs. Berry (4) reported improved legibility in the handwriting of students in grades 1–8 when they received instruction in italic writing. Simplifications of both letter forms and handwriting style were also introduced and studied in Russia. Articles (79, 14) describing the simplified approach reported that achievement of first-form pupils was accelerated in both reading and writing.

Another approach to the study of letter forms has been through the use of factor analysis to determine commonalities among letter shapes. In a series of studies Kuennapas (45, 46) identified three factors in the formation of nine capital letters; in a more extensive study, Kuennapas and Janson (47) factor analyzed the twenty-eight lowercase letters of the Swedish alphabet. Two factors—called vertical linearity (as in "t") and roundness (as in "o")—accounted for a considerable portion of similarity among many lowercase letter forms. Although considerably more research of this type is needed, the identification of commonalities of letter forms could influence the sequence of instruction when letter forms are initially taught. In other words, the line movements that are common to various letters could be taught first, followed by instruction in the specific letter forms. This procedure, common to many instructional programs, is currently based on logic rather than on objective evidence of commonalities.

Future research ought to focus upon both the identification of the most desirable modifications in letter forms and the impact of such modifications upon handwriting instruction and the quality of adults' writing.

Instructional Techniques and Correlates of Skill Development in Handwriting

In this section research directed toward clarification of how children ought to learn to write is considered. Herrick and Okada (39:30) probably had a similar concern when they posed this question:

> *Most of the teaching aids deal with tracing, copying, and manually guiding the child in the act of learning how to write. There also seems to be growing evidence that growth in writing skill plateaus rapidly from the third grade on. Could it be that too much attention is paid to tracing, copying, and manually guiding, and too little to the perceptual processes of the learner himself as he goes beyond form to normative and aspirational percepts of his own writing?*

Unfortunately, the literature reveals little concern for the role of tracing, copying, or manual guidance in learning to write. Even outside the area of handwriting, little research has been done. Birch and Lefford (6), for example, repeated an observation by Smith and Smith (75) that *"One is struck by the absence of any systematic account of perceptually-guided motion within our general behavior theories"* (6:1). Although Birch and Lefford did not deal directly with handwriting, a brief discussion of their research may indicate a fruitful direction for handwriting research. They proposed a model in which the development of voluntary motor control is a function of visual-perceptual competence on the one hand and improved inter-sensory liaison between vision and body sensations on the other. To test the model, they examined the ability of children, ages 5 to 11, to draw geometric forms from visual models. Four figures (upright isosceles triangle, inverted isosceles triangle, diamond with an elongated vertical axis, and equiangular diamond) were presented to the children under six conditions of varying stimulus support:

> *1. Tracing task. (The child placed his response sheet over the models and traced the lines.)*
> *2. Connecting-dots condition. (Dots were placed at the points where lines intersected and ended on the models; on the Ss' response sheet only the dots were printed in the same positions as on the models.)*
>
> *3. Dot grid condition. (The models were superimposed on a grid of dots; the child's response page contained only the dot grid.)*
> *4. Line grid conditions. (Similar procedures as with the dot grid except that the grid was composed of straight, intersecting lines.)*
>
> *5. Freehand drawing from model with dot emphasis. (The models were the same as used in the connecting-dots condition; the S copied the models on a blank page.)*
>
> *6. Freehand drawing on a blank page. (The S copied the models on a blank page.)*

The conditions were markedly different in difficulty for the young children; as the Ss approached age 11, they were able to perform competently under all conditions although significant differences in accuracy level existed between tracing and freehand drawings even at age 11. In general, as stimulus support was removed, younger children did poorer. That is, the younger children had the highest accuracy scores on tracing, followed closely by the dot-connecting task. The lowest scores were obtained on the freehand drawing on a blank page, with the two grid-background conditions intermediate in performance level. (The freehand drawing condition from the models with dot emphasis was similar in difficulty to the freehand drawing with the standard models.) The freehand drawing clearly remained the most difficult condition for all ages, and improvement with age varied with the model being copied.

The implications of the Birch and Lefford study for handwriting instruction seem clear. Since tracing was the easiest task for children, especially those beginning elementary school, and freehand drawing the most difficult, perhaps instructional programs in handwriting should teach the letter forms through tracing, not through copying the forms on a blank page. A variation from the tracing procedure could be found in the connecting-dots condition in which a dot would be placed at every end point and at every intersection of lines. Even older children might well be taught by tracing when they make the transition to cursive writing or when they receive remedial instruction on particular letter forms. Herrick and Okada (39), however, discovered in their national survey of instructional practices that copying was the teachers' preferred teaching device while tracing was listed as the third preference. As the investigators pointed out, these practices do not result from research findings so much as from traditional instructional procedures.

An additional consideration is that growth in the ability to copy geometric forms was found by Townsend (80) to correlate significantly more highly with mental age than with chronological age. Although Birch and Lefford (6) made no provision for the study of varied IQ levels in their study, the Ss were described as "bright normal" with a median IQ of 115 and with no child's IQ below 90. It seems reasonable, therefore, to conclude that those Ss were bright enough to perform the copying task and that they would have excelled in it if copying were actually superior to tracing as an instructional procedure.

Herrick and Okada commented also on the need for research regarding learners' perceptions of their own writing. Harris and Herrick (32), drawing their discussion of this problem from a more extensive document (38), investigated the ability of bright, average, and slow children to form "normative" and "aspiration" models for their handwriting. Normative perception has to do with a child's ability to identify a sample of handwriting as like his own; aspirational perception has to do with the model that a child sets as a goal for himself. In a task in which fifth and sixth graders were asked to rank handwriting samples on a 5-point scale, bright and average students did significantly better than slow students according to adult criteria, although the judgments about correct ordering by all pupils were not identical with the judgments by adults. All children, however, were able to identify an aspirational model as at least equal to, or better than, handwriting like their own; in fact, their choices of an aspirational model appeared to be more consistent than those involving identification of a normative model. Even over a period of time children did not improve in their normative perceptions of their own handwriting. The conclusion of the investigators was that instructional emphasis should be placed on the development of normative and aspirational models with less emphasis on letter formation drill. An implication that may be drawn from the study is that the normative perception—the ability of a child to pick out a handwriting sample like his own—must be well established before a child can be expected to modify his own handwriting as a result of comparisons to an aspirational model. In other words, if a child does not have a clear perception of his own handwriting, use of an aspirational model (the standard instructional letter form, for example) will be ineffective in producing change. Development of normative perception should take place in the primary grades along with the use of an aspirational model.

Studies of correlates of handwriting success have focused mainly on the relationship of intelligence and handwriting legibility. Love (53) compared mentally retarded and normal Ss of prevocational and vocational ages on the basis of handwriting quality, speed, and use. Handwriting samples of the retarded Ss were rated significantly higher in quality on the Ayres scale than those of normal Ss; but the retardates wrote significantly more slowly. Conclusions to be drawn from the study are severely limited by the fact that (a) only twenty-five Ss were studied, and (b) two age groups (12–15 and over 15-years-old) were considered together for each intelligence group. Although Lewis (48) was not concerned with the same age groups nor with extremes in IQ, he concluded that at the first-grade level mental maturity was more closely related to writing errors before than after instruction. In other words, after instruction IQ was no longer significantly related to writing performance. Lindell (51) investigated factors related to the handwriting ability of eighty Swedish Ss in grades 3–6 and found that intelligence was only moderately related. Yet the most important factor proved to be ability to use the spoken language, which would appear to be not unrelated to intelligence. Finally, Perron and Mignard (64) evaluated the handwriting of 243 French mental retardates and found that achievement in handwriting was higher than expected on the basis of mental age or achievement in other school subjects. The investigators concluded, however, that the relative importance of intelligence depends upon the stage of development of the child. The data from these studies do not clearly indicate that special instructional programs in handwriting are warranted for mentally retarded children. Yet, various adapted instructional programs, which are discussed in a later section, are suggested for these children and may in fact be necessary only during initial learning of handwriting skills.

Effects of Various Body Part Positions on Writing Performance

Regarding the effect of body position upon handwriting, Lurcat (54) found that posture was the most important factor affecting the slant of the handwritten line. Only two of her sixty French Ss, who ranged from 4- to 7-years-old, were able to write lines that were parallel to the edge when given unlined paper. More research in this area is obviously needed.

Callewaert (11) has described in detail a modified handwriting grip that he claims is physiologically superior to the traditional grip because muscles are more relaxed so there is less cramping. The recommendation is that instead of holding the barrel of the pen between the thumb and index finger as in the traditional grip, the shaft be placed between the middle and index fingers and the wrist be more sharply canted. Callewaert also described a training program for kindergarten children that would prepare them for handwriting instruction in the first grade. No evidence was offered—other than logic and a knowledge of physiology—for the superiority of the modified grip.

The use of Callewaert's modified grip was, however, evaluated by Otto and others (62). Twenty right-handed college students were asked to copy unfamiliar material during two 30-minute sessions, using the traditional grip and the modified grip with a 1-week interval between sessions. The Ss favored the traditional grip in terms of comfort, and the legibility ratings of the words produced with the traditional grip were higher. However, because the Ss did adapt quite readily to the modified grip, the authors concluded that further research is needed, especially with objective measures of muscular tension.

Comparisons of left- and right-handed Ss' handwriting performance have been made in a number of studies. Groff (29) matched, by sex and grade, 615 right-handed children with the same number of left-handed children in grades 4, 5, and 6. Using the Ayres legibility scale and five judges, he found only one significant difference between legibility ratings: right-handed girls in sixth grade had higher mean ratings than left-handed girls in the same grade. Groff concluded that the data did not demonstrate a general superiority of right-handed writers. In another article, Groff (27) compared right- and left-handed writers in grades 4–6 on speed of handwriting. Although girls were significantly faster than boys, no differences in speed were found with respect to hand preference.

Reed and Smith (70) also examined the speed and quality of left- and right-handed children's writing. Working with 10-, 12-, and 14-year-olds, they matched right-handed Ss to left-handed Ss with respect to age and sex. No significant differences due to handedness were found on speed of writing, as measured in a timed test of simple repetitive writing or a timed test of copying a prose package. When eight judges were asked to choose the better sample from each pair of Ss, again no significant differences were found. Finally, no significant differences in either speed or quality were found between the two groups on a fatigue test which involved an extended period of copying a prose passage. Lewis (48) analyzed the ability of first graders to copy the manuscript alphabet and found that left-handed children made more errors than right-handed children before formal handwriting instruction. After instruction, however, no significant differences were found in the total number of errors, although left-handed subjects made significantly more reversal and inversion errors.

Taken together, the Groff, Reed and Smith, and Lewis studies do not support the commonly held notion that right-handed writers write more rapidly and legibly than left-handed writers. Nor do they support the findings of an earlier study (31) in which right-handed fifth and sixth graders, who were matched to left-handed writers by IQ, age, grade, and sex, were superior on measures of handwriting speed and quality. Perhaps teachers are heeding fairly recent but repeated admonitions to make some special provisions for left-handed writers. This would amount to a positive explanation for an apparent improvement in left-handers' writing performances between the 1930's and the 1960's.

On the basis of classroom observations in grades 5–8, Enstrom (15) identified fifteen techniques of adjusting positions for handwriting used by left-handed writers. These, however, fell into two general categories—one where the writing hand is kept below the line and the other where the "hook" wrist is employed. Handwriting samples from 1,103 Ss were judged for quality, speed, neatness, and "healthful body positions." The position given the highest overall rating was not the one commonly taught as an adjustment for left-handed students. Drawings of adapted writing positions presented in the article should be very helpful to teachers in providing instruction for left-handed writers. Other articles, such as one by Croutch (13), also present suggestions about the correct position for the body and the paper for both left- and right-handed writers.

Effect of Speed and Stress on the Handwriting Product

While the writing speed of left- and right-handed writers has been compared, the general effect of speed upon legibility apparently has not been examined in the past 10 years. There was, however, an attempt to define new standards for measuring speed. Groff (30) devised new speed norms for grades 4, 5, and 6. He obtained writing samples from 4,834 pupils and compared them with the older Ayres speed norms. Instead of having the children copy a standard sample until their teachers were satisfied that they were familiar with the words in it, as was done by Ayres in devising his norms, the children read the passage until they could recognize all the words. Then the speed of writing was measured in terms of time consumed in copying the passage for the first time. Groff concluded that his new norms, with rates that are considerably slower than those reported by Ayres, are more defensible than the long accepted Ayres norms because the procedures followed in producing them were more realistic. The authors concur with this conclusion.

Rarick and Harris (69) investigated the writing speed of bright, average, and slow sixth-graders under four different writing conditions: normal, fast, best, and relaxed. The bright and average group produced their fastest writing sample under the condition calling for fast writing; usually the slowest writing was produced under the condition calling for "best" writing. Although sex was found to be a highly significant factor, legibility was also a function of the condition of writing. That is, high legibility tended to occur with a slow rate. The correspondence is not perfect, however, in that generally superior writers tended to produce the most legible samples under all conditions. Degree of stress was also measured in terms of galvanic skin response, muscular tension, and fine motor control. In general, for the bright and average groups the greatest stress reactions, which tended to occur in the fast writing condition, were accompanied by lower legibility. Conversely, the least stress was observed during relaxed writing, although legibility was not universally improved. No observable differences in affective states were noted among the different IQ groups. However, when average and bright groups were combined, greater stability in affective measures and in legibility was observed across writing conditions than in the groups of slow learning children.

Indications of stress are also obtained from measures of finger pressure exerted on the barrel of the writing instrument, point pressure of the writing instrument on the writing surface, and pressure of the hand resting on the writing surface. Herrick and Otto (41) measured pressure of the fingers on the barrel of the writing instrument and found wide individual variation. In general, however, high point pressure accompanied high grip pressure and low point pressure accompanied low grip pressure. The study represents a step forward in terms of the instrumentation employed, but much additional study of the relationship of stress and handwriting performance is needed before instructional implications will be clear.

Effect of Instructional Sequences on Development of Handwriting Skills

The research pertaining to handwriting instruction is best introduced by a brief consideration of certain surveys of current practices in handwriting instruction.

Surveys of Instructional Practices. Herrick's extensive inquiry (34) into the instructional practices advocated by nineteen commercial systems of handwriting instruction revealed many common practices, but a few differences were also identified. No general consensus regarding sequence of instruction was observed although, in general, letters classified by similar shapes—e.g., straight line, slant line, or circle—were introduced together. The greatest variations in letter form models advocated by different systems occurred in the uppercase cursive letters and in numbers; there were fewer differences among lowercase cursive letter forms. Photographic reproductions of the letter forms used by the nineteen companies surveyed were presented for comparison by Herrick and Otto (40). Noble (58) reported a survey of sixteen commercial systems of handwriting instruction.

A number of surveys of instructional practices in the schools were also reported. For example, a national survey of six hundred school systems made in 1962 (39) revealed that 98 percent of the teachers reported that they do indeed teach handwriting. Most schools reported a separate handwriting class-period of approximately 15–20 minutes five times a week in grades 1–4 and three times a week in grades 5–8. Most schools also reported that they incorporate handwriting instruction in a meaningful context in conjunction with subject matter areas. Only one-fifth of the schools reported attempts to individualize handwriting instruction. King's survey (44) of 680 school systems in four midwestern states revealed that of the fourteen systems of handwriting instruction being used, two publishers supplied 89 percent of the instructional material.

Some commentators on handwriting instruction (e.g., 16) have claimed that handwriting instruction—and consequently the quality of handwriting—has declined in recent years. Erlebacher and Herrick (19), however, found no substantial differences in the quality of handwriting of sixth-grade children in 1912 and 1959 according to Ayres scale ratings. Anderson (3) compared handwriting samples from nine hundred children in England, Scotland, and the United States, respectively, to determine whether handwriting quality would excel where a "traditional" instructional approach is employed, as in Scotland. Drawing subjects from 7-, 11-, and 14-year-old age groups, he found that the English and Scottish children tended to lose some initial advantage in quality as age increased. In fact, the handwriting of the Scottish sample tended to deteriorate after age 11, while the mean ratings of the American sample tended to increase with age. It appears, therefore, that "traditional" instruction, operationally defined as that used in Scotland today and in the United States early in the present century, does not necessarily produce better handwriting than modern instructional programs.

Evaluation of Instructional Programs. Extensive comparisons of the effects of different instructional systems upon pupils' handwriting achievement have not been made. Most research pertaining to instruction has been tied to the development of new instructional approaches. Existing programs, then, have been evaluated only for the purpose of comparison with a new approach.

The effects of individualized procedures in teaching handwriting were examined in two studies. Tagatz and others (77) compared the effects of three approaches to handwriting instruction with eighty-eight third-grade and seventy-five fourth-grade pupils who were stratified by sex and previous achievement/ability level. Although the duration of the study was only 9 weeks, two approaches that stressed individualization produced greater gains in legibility in third graders than a formal instructional approach, which involved the continued use of a commercial system. Fewer significant treatment effects were obtained with the fourth-grade sample, and no one treatment produced superior results. One implication drawn from the study was that individualized approaches which emphasize the diagnosis and remediation of specific difficulties may best be employed in the primary grades; however, comparison of different approaches to handwriting instruction at various grade levels needs to be made for a longer period of time than 9 weeks before any definitive conclusions can be drawn.

Boyle (8) compared the handwriting achievement of 312 children in grades 4–6 who were enrolled in control and experimental classes. In the experimental treatment instruction was planned on the basis of diagnostic testing. Although no differences between experimental and control Ss were obtained on gross measures of handwriting quality (i.e., ratings of handwriting samples on legibility scales), Ss in the experimental group eliminated significantly more errors than the control group in the size, slant, and formation of letters. The results of this study again give evidence that pupils can improve their handwriting if they are taught by an individualized approach. One can only speculate that the differences between experimental and control Ss might have been greater had the experimental treatment been administered to pupils in the primary grades before handwriting habits had become established.

Some earlier studies, too, seem to support the superiority of an individualized, diagnostic program of instruction. Newland (57), for example, found that four types of difficulties in letter formation caused over half of all illegibilities; he also noted that four letters constituted about 45 percent of all illegibilities in the handwriting of elementary, high school, and adult Ss. Newland, therefore, concluded that both developmental and corrective instruction could be highly specific. Application of Newland's findings would involve the diagnosis and remediation of problems peculiar to each child, the assumption being that by correcting specific illegibilities the general quality and speed of the handwriting would be improved.

Only one instructional approach developed since 1960 was described extensively. The unique feature of that approach is that there is an attempt to coordinate handwriting instruction with children's perceptual development. Furner (22, 23, 24), in a series of articles, contended that since handwriting is a type of perceptual-motor learning, methods of instruction should reflect an awareness of perceptual development in children. Basing her discussion on the work of Hebb, Furner argued that perceptual development is a learned process which must necessarily involve movement. In her opinion, however, existing instructional programs in handwriting focus only on motor development, treating handwriting exclusively as a motor process, not as a perceptual process as well.

The first step in Furner's instructional program involves the visual component: the child sees the finished product (i.e., the letter) as well as the formational process. Then the child is guided in verbalizing the formational process involved in production of a written symbol—i.e., describing the starting point, direction, size, and stopping point. This auditory process ostensibly reinforces the previous step of visualization. Since the letters are introduced in groups according to commonalities in shape (e.g., straight lines, circles), the child is also asked to state these generalizations. Throughout the instructional process the habit of self-evaluation and correction is established as an aid to perceptual development and as a basis for acquisition of the necessary motor skills. These procedures are based on Harris and Herrick's (32) recommendation that normative and aspirational models be developed from the beginning of handwriting instruction.

Furner evaluated her instructional approach in grades 1–3 by comparing test results from her program in four experimental schools with the results obtained from use of commercial programs in seven control schools. The study was continued for 3 years, with a new group of first graders introduced each year. Thus, at the end of the study data had been gathered on a group of third graders over a 3-year period, second graders for 2 years, and first graders for 1 year. The quality and speed of handwriting of experimental and control Ss were compared at midyear and at the end of each of the 3 school years, and formational errors were also compared. As for the quality of handwriting, experimental Ss were significantly superior to control Ss in grades 1 and 2. In grade 3, however, control Ss were superior in both the February and June testings. Furner suggested that the shift might have been due to teachers' unfamiliarity with the experimental approach, which was responsible for a delay in the transition from manuscript to cursive writing. Results with regard to speed were mixed: the control group scored significantly better only in June of first grade, February of second grade, and February of third grade; at other testings no significant differences were found. Formational errors in randomly selected Ss were observed at the end of each of the 3 years of the study; the experimental Ss consistently made fewer formational errors than the control Ss.

In light of the significantly fewer errors in the formational processes of writing letters and the significantly higher quality scores in grades 1 and 2 for the experimental Ss, Furner's instructional program does appear to have certain advantages over the control systems. The approach appears to be theoretically sound if the premise is accepted that handwriting is a learned perceptual-motor process and that, therefore, instructional programs ought to emphasize both the perceptual and the motor aspects of skill development.

Birch and Lefford (6) have shown that intersensory integration, particularly the ability to integrate visual and kinesthetic information, is highly and significantly correlated with the ability to copy a visual model in 5-, 6-, and 7-year-olds. Thus, teaching the child to integrate different sensory modes could lead to increased handwriting skill. Furner did not, however, include tracing as an instructional technique. Instead, she emphasized the notion that a mental image should be created and the letter copied from that. In light of the findings of Birch and Lefford, tracing might advantageously be inserted as a step prior to forming a mental image. Kinesthetic reinforcement would thus be given to the visual perception of a letter in creating the mental image Furner suggests.

Other research pertaining to instruction in handwriting has been concerned with suitable programs for special education classes. Suggestions for teaching handwriting to the slow learner as well as to the brain damaged child have been offered by Linn (52), Harrison (33), Enstrom (17), and McElravy (55). Green (25) devised an instructional program for twenty-four brain damaged children and concluded, in line with most writers on the subject, that a kinesthetic approach is best. Parres (63), also investigating handwriting instruction in educable mentally retarded classes, concluded that the handwriting habits of some children, especially those who are younger and more handicapped, are disrupted by tasks which require crossing the body midline. His recommendation was that early handwriting take place only on the side of the preferred hand.

Very little has been written concerning instructional programs in handwriting for teachers. King (44), in a survey of 680 school systems, found that only 9 percent required some kind of training in handwriting for elementary teachers. Plattor (66) studied various types of preservice training programs for teachers: two groups of preservice teachers were given specially designed programmed instruction in teaching handwriting; one group had conventional classroom instruction in the teaching of handwriting; and one group had no training. The Ss who received programmed instruction scored significantly higher than the Ss who received conventional classroom instruction on a test of their knowledge of teaching handwriting; programmed instruction also proved to be as effective as conventional instruction in improving the S's own handwriting (i.e., in improving his ability to serve as a model for children). The groups that received handwriting instruction were significantly higher on both types of measures than the control group that received no instruction. On the basis of this study it appears that some kind of instruction in teaching handwriting is beneficial in preparing teachers. Future research should be designed to investigate whether or not teacher training ultimately has an effect on the handwriting of children who receive instruction.

Manuscript Versus Cursive Writing. The questions of which style of handwriting should be used initially in handwriting instruction and whether an initial style should be retained have been debated at great length. The purpose here is not to state another opinion but to examine the literature to see if any viable conclusions can be drawn from it. Unfortunately, though, more opinions than research have been published within the last 10 years.

Herrick (35) presented the arguments for each style of writing and summarized the research prior to 1960 pertaining to each. His conclusion favored the retention of manuscript writing beyond the primary grades; but he also stated that the question is probably not important to most adults, who should be allowed to decide on their own style. Hildreth (42) summarized earlier research on the question and concluded that wide individual differences existed in students at the upper-grade levels. Speed of handwriting seemed to be more closely related to instruction, practice, and the characteristics of the person than to his style of writing. Andersen (2), who examined earlier research on the two styles in terms of the issues of legibility, speed, and ease of learning, concluded that manuscript writing not only is more legible than cursive but also can be written at least as fast and learned more easily by children and adults. He suggested that longitudinal research is needed to provide definitive evidence of the superiority of the manuscript style.

At least six research studies relevant to the manuscript–cursive issue were reported in the 1960's. Three of them compared the number of errors made while children used each style, two evaluated speed and pressure with each style, and one examined the legibility of adult writing with each style.

Meis (56) had 3,750 beginning school children copy ten words in both manuscript and cursive styles. Significantly fewer errors were made on each word when the words were printed in manuscript style. The immediate conclusions to be drawn from this study are, however, limited by the fact that German script rather than English was involved.

Byers (10) used the same dictated paragraph to test two groups of third graders who had received 8 months of instruction in cursive writing. One group wrote the paragraph in cursive first and in manuscript 10 days later; the procedure was reversed for the other group. Although the two groups were supposedly equated, one group proved to make more spelling errors in writing in both styles. For individual children, there was no significant difference in spelling errors between the two styles. In analyzing the kinds of spelling errors made, the following tendencies were observed: More transpositions and omissions of letters occurred when the manuscript style was used; more substitutions of letters and omissions of entire words were made with cursive writing. The data, therefore, do not favor either style.

Bolen (7) randomly assigned 312 third graders to twelve classrooms in ten schools, six of which made the transition to cursive writing at the beginning of the school year, and six of which made a delayed transition after the study was completed. Prior to the study both groups were given an intelligence test and a standardized spelling test. Instructional procedures in spelling during the school year were equated in terms of method and time for both groups. At the end of the year, before the manuscript group made the transition to cursive writing, the spelling test was readministered and gain scores for the two groups were compared; no significant differences were found between the group using manuscript writing and the group that had made the transition to cursive writing.

None of the above studies offer conclusive evidence for the superiority of either style in terms of the number of errors made. Although the style of writing may affect the ability of young children to copy words, it apparently does not make any difference in terms of spelling abiltiy once children have progressed beyond initial instruction in handwriting.

Weinert, Simons, and Essig (81) compared the speed and pressure of handwriting of 672 German children after 4 years of school. Some of the children had been taught manuscript writing initially and some of them had used cursive writing from the beginning. They found that those who had received initial instruction in manuscript style wrote faster and had less trouble with cramping. Sokolova (76), however, found in Russia that experimental classes that used cursive writing from the beginning of instruction increased their speed of writing 50 percent and almost 30 percent over control classes (using manuscript) in the first and second classes, respectively. In the third and fourth classes experimental and control groups were not significantly different in speed. By the end of the fourth class and in the fifth class, however, experimental Ss wrote more rapidly than control Ss. This type of longitudinal data should be gathered in English-speaking countries to compare the merits of initial instruction in cursive and in manuscript writing.

Templin (78) conducted a study to determine whether manuscript or cursive writing was more legible when used by adults 10 years after high school graduation. Subjects who had been trained only in manuscript writing in elementary school achieved a higher mean legibility score than either the group that had received only cursive instruction or the group that had received manuscript instruction followed by a transition to cursive. In fact, the latter group of Ss, those who had first learned manuscript and then made the transition to cursive, had the lowest mean legibility rating among the three instructional groups. However, Templin concluded that the quality of handwriting was also influenced by characteristics of the writer in addition to the type of early instructional program. Another aspect of the study focused upon the use made of each style. The Ss had been asked for relevant information before they knew they were part of a handwriting study. Males, regardless of early training, used manuscript more often than cursive writing; females, on the other hand, used cursive writing more often than manuscript.

The latter finding appears to indicate that style of handwriting is tied to the identification of sex roles. Males may have felt that the manuscript style appears more business-like. Groff (28) surveyed the personnel directors of 115 large, successful corporations to determine if a preference existed for one style in the handwriting of employees. On the basis of 80 percent response to the survey, Groff found that 53 percent of the personnel directors said either style was acceptable and 33 percent indicated a preference for manuscript writing. Of those who did state a preference, then, the majority preferred manuscript writing—a finding that is in line with Templin's discovery that males, who are more likely than females to be in business occupations, preferred manuscript writing.

For practical purposes, there seems to be no need to change from manuscript to cursive writing if manuscript writing is introduced first. Yet definitive data, to show that the introduction of manuscript writing does indeed produce superior results, have not yet been reported.

Transition Time. Other studies have been concerned with the timing of the transition from manuscript to cursive writing. Groff (26) surveyed directors of elementary education in metropolitan areas and concluded that time of transition is determined by tradition rather than by research results. Otto and Rarick (61) studied the handwriting, spelling, and reading performance of fourth and sixth graders who had used the same instructional materials but had made the transition at the four most common times—first half of second grade, second half of second grade, first half of third grade, and second half of third grade—to determine if one particular transition time would be associated with superior performance. Transition time was significantly related only to handwriting performance: a late transition was associated with rapid writing and an early transition was associated with legible writing. The conclusion drawn by the investigators was that transition time is not as important as the nature of the instructional program in terms of the effect upon later achievement in handwriting, spelling, and reading. Further investigations of transition time would probably serve only to underscore this conclusion.

Nature of Handwriting Instruments and Writing Surfaces That Facilitate Learning to Write

Apparently there have been no studies of writing surfaces in the past decade, but a few studies of writing instruments were reported. Herrick's
studies are most notable in this area. In one artic
(36) he reported that children prefer adult pencil
over beginner pencils, which have not been show
objectively to be better for beginning writer
In another study (9) the general conclusion wa
reached that the use of ballpoint pens is condu
cive to deterioration of handwriting quality.
group of industrial design students (82) reporte
the development of an especially devised instru
ment for handicapped people who are unable t
use conventional writing tools.

Obviously, the research in this area is too scant
to permit any conclusions. More work is neede
to determine what, if any, modifications in wri
ing instruments are associated with easier, faste
and/or more legible writing.

Development of New Handwriting Scale and Evaluative Materials, and Means for Helping Teachers Make Better Use of Existing Scales

Two attempts have been made between 196
and 1970 to find new methods of scaling han
writing samples for legibility. Herrick and Erl
bacher (37) rejected the idea of a single sca
with one sample for each quality level by whic
to judge handwriting samples when several d
mensions exist. They described a series of studi
which led to the scaling of six hundred samples-
two hundred each from grades 4, 5, and 6—i
terms of size, slant, and legibility. Because of th
fine distinctions with such a procedure, the mast
list of six hundred scaled samples is limited
research use. In the future perhaps scales shou
be built on the information obtained from fact
analysis of handwriting characteristics. For i
stance, Fischer (21) analyzed twenty-five chara
teristics of handwriting in Germany and foun
five factors, two of which were size and slant.

Bezzi (5) also used separate scales for diffe
ent grade levels. His series of scales for grad
1–3 are unique in that they are intended for u
with manuscript writing, whereas other scal
are to be used for judging cursive writing sa
ples only. His 5-point scales were devised on t
basis of analysis of 7,212 handwriting specime
from 130 schools throughout the United States.

Feldt (20), rather than devising a new sca
studied the reliability of a scale used as part
a commercial instructional program in handwr
ing with the intent of improving the use of sc
ing in the schools. Using three samples taken ov
a period of time from each of thirty first grade
and twenty-nine second graders, three judg
rated them independently on the commercial sca
After a thorough analysis of the data in vario
combinations, Feldt suggested that in order
raise reliability of ratings, scores from sever
independent testing sessions be averaged and mo

complete instructions for using the scale (including practice samples) be provided. Feldt's suggestions seem to have direct application for teachers who use commercial scales and for publishers who provide scales.

Rondinella (71) found that teachers not trained in the use of scales were subjective in their ratings of handwriting samples. Samples from 239 children in grades 4, 5, and 6 were rated by 210 intermediate teachers. The teachers' ratings were then compared to those of experienced raters using published scales. Although the correlation of judgments by the two groups was significant, it was low (.41); inconsistencies among teachers' ratings of particular samples were also great. When teachers were asked to state what criteria they used in making their judgments, they named fourteen, only five of which are considered with the widely used Freeman Scale. Letter formation was the most frequently named criterion measure. Rondinella concluded that teachers should be trained in using scales if they are to make use of them in the diagnosis and treatment of the handwriting difficulties of their students.

Otto, Askov, and Cooper (60), however, found that once teachers are familiar with using a scale, they can continue to rate samples without the use of a scale. Three judges independently rated the handwriting samples of 240 fourth- and sixth-grade students using two published scales. Then the judges assigned legibility ratings without using a scale. Since inter-judge reliability coefficients were about as high without the use of the scales as with them, the conclusion was drawn that experience with scales allows teachers to make as accurate judgments about legibility without the scales as with them. Thus, the suggestion was that teachers can discontinue the use of scales once they have internalized some basis for making judgments regarding legibility.

Finally, Craig (12) studied the handwriting factors affecting both teachers' ratings of legibility and pupils' ability to read the passages written by their grade peers. Three samples were obtained from each sixth grader in three elementary schools, making a total of 240 samples. The investigator assumed that by sixth grade any difficulty in reading a sample would be the result of illegibility of the sample, not of a basic inability to read, an assumption that seems questionable at best. The main finding was that the handwriting factors that were significantly correlated with teachers' ratings also were correlated with the oral readings of peers. Apparently, teachers employed factors pertaining to readability in making their judgments about legibility. The following specifics, rather than a more global quality of "readability," were found to be important: height of small letters, uniformity of small letters, uniformity of tall and capital letters, and alignment.

Scales that are developed for research purposes must permit fine discriminations among samples; thus, many quality levels must be provided. When such scales are brought into the classroom, however, ratings by individuals tend to have low reliability. Reliability can, of course, be increased by obtaining ratings from several judges, but such a procedure tends to be cumbersome in application. Pragmatically, an internalized scale developed through training and experience will probably be most useful in the classroom. Future research, then, ought to focus on means for bringing about such internalization and on the identification of specific factors that are critical in judging legibility.

A Final Word

While progress has clearly been made in the past decade, it seems equally clear that the progress has been neither rapid nor dramatic. Yet, one might wonder why, in a world beset with horrendous problems, there is any concern at all with a subject so mundane as handwriting. There are those who say it is, after all, nothing more than a highly personalized means for expression, that a typewriter or a printing machine or a computer can do the job much better and faster. Perhaps, but in a frighteningly impersonal world a highly personal means for communication may have much to recommend it. We think so. Yet the fact remains that handwriting cannot become so personal that it is completely idiosyncratic and illegible to all but the writer. Earl Wilson[1] seems to have caught a glimpse of where that road ends:

> *A schoolboy brought home his card*
> *with a teacher's notation to his parents:*
> *"Your son's handwriting is so bad that*
> *we don't know if he can spell."*

So we would like to see the art advance. Who can tell what the next decade will bring?

FOOTNOTE

1. Wilson, Earl, "It Happened Last Night," *Wisconsin State Journal*, Madison, Wisconsin, March 18, 1970, p. 16.

REFERENCES

1. Andersen, Dan W., "Handwriting Research: Movement and Quality," in Horn, Thomas A. (ed.), *Research on Handwriting and Spelling*, National Council of Teachers of English, Champaign, Illinois, 1966, pp. 9–17.
2. Andersen, Dan W., "Handwriting Research: Style and Practice," in Horn, Thomas A. (ed.), *Research on Handwriting and Spelling*, National Council of Teachers of English, Champaign, Illinois, 1966, pp. 18–28.
3. Anderson, Irving H., *Comparisons of the Reading and Spelling Achievement, and Quality of Handwriting of Groups of English, Scottish, and American Children*, United States Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1903, University of Michigan, Ann Arbor, 1963.

4. Berry, Winifred, "Italic Writing," *Education Digest*, 26:50–51, 1961.
5. Bezzi, R., "A Standardized Manuscript Scale for Grades 1, 2, and 3," *Journal of Educational Research*, 55:339–340, 1962.
6. Birch, H. G.; Lefford, A., "Visual Differentiation, Inter-sensory Integration, and Voluntary Motor Control," *Monographs of the Society for Research in Child Development* 32: (whole no. 110) 2, 1967.
7. Bolen, John E., "A Study of Manuscript Writing and Spelling Achievement in the Third Grade," *Dissertation Abstracts*, 25: (no. 11) 6422, 1965.
8. Boyle, Sister Mary Charles, O.P., "An Experimental Study of a Diagnostic-Remedial Program in Handwriting in the Fourth, Fifth, and Sixth Grades," *Dissertation Abstracts*, 24: (no. 2) 642, 1963.
9. *Bulletin of the International Bureau of Education*, 37: (no. 147) Geneva, second quarter, 1963.
10. Byers, Loretta, "The Relationship of Manuscript and Cursive Handwriting to Accuracy in Spelling," *Journal of Educatioinal Research*, 57:87–89, October 1963.
11. Callewaert, H., "For Easy and Legible Handwriting," in Herrick, Virgil E. (ed.), *New Horizons for Research in Handwriting*, University of Wisconsin Press, Madison, 1963, pp. 39–52.
12. Craig, Myrtle C., "An Analysis of the Relationships Between the Ease of Reading Sixth Grade Handwritten Papers by Peers and Teacher Evaluation of the Handwritten Papers with Selected Handwriting Factors," *Dissertation Abstracts*, 27: (no. 2) 325-A, 1966.
13. Croutch, Ben, "Handwriting and Correct Posture," *Academic Therapy*, 4:283–284, Summer 1969.
14. deVette, Robert O., "Notes and News: Russian Penmanship Simplified," *Modern Language Journal*, 50: 220, 1966.
15. Enstrom, Erick A., "The Relative Efficiency of the Various Approaches to Writing with the Left Hand," *Journal of Educational Research*, 55:573–77, August 1962.
16. Enstrom, Erick A., "The Decline in Handwriting," *Elementary School Journal*, 66:22–27, October 1965.
17. Enstrom, Erick A., "Out of the Classroom; Handwriting for the Retarded," *Exceptional Children*, 32:385–388, 1966.
18. Epstein, Lawrence; Hartford, Huntington; Tumarkin, Irving, "The Relationship of Certain Letter Form Variants in the Handwriting of Female Subjects to Their Education, IQ, and Age," *Journal of Experimental Education*, 29:385–392, June 1961.
19. Erlebacher, Adrienne; Herrick, Virgil E., "Quality of Handwriting Today and Yesterday," *Elementary School Journal*, 62:89–92, 1961.
20. Feldt, Leonard S., "The Reliability of Measures of Handwriting Quality," *Journal of Educational Psychology*, 53:288–292, 1962.
21. Fischer, Gerhard, "Zur Faktoriellen Sturktur der Handschrift" (Factorial Structure of Handwriting), *Zeitschrift fur Experimentelle und Angewandte Psychologie*, 11:254–280, 1964.
22. Furner, Beatrice A., "The Perceptual–Motor Nature of Learning in Handwriting," *Elementary English*, 46: 886–894, 1969.
23. Furner, Beatrice A., "Recommended Instructional Procedures in a Method Emphasizing the Perceptual–Motor Nature of Learning in Handwriting," *Elementary English*, 46:1021–1030, 1969.
24. Furner, Beatrice A., "An Analysis of the Effectiveness of a Program of Instruction Emphasizing the Perceptual–Motor Nature of Learning in Handwriting," *Elementary English*, 47:61–69, 1970.
25. Green, Mary I., "An Introductory Study of Teaching Handwriting to the Brain Injured Child," *Exceptional Children*, 34:44–45, 1967.
26. Groff, Patrick J., "From Manuscript to Cursive—Why?" *Elementary School Journal*, 61:97–101, 1960.
27. Groff, Patrick J., "Who Writes Faster?" *Education*, 83:367–369, 1963.
28. Groff, Patrick J., "Preference for Handwriting Style by Big Business," *Elementary English*, 41:863–864, 868, December 1964.
29. Groff, Patrick J., "Who Are the Better Writers—the Left Handed or the Right Handed?" *Elementary School Journal*, 65:92–96, November 1964.
30. Groff, Patrick J., "New Speeds in Handwriting," in Otto, W.; Koenke, K. (eds.), *Remedial Teaching: Research and Comment*, Houghton Mifflin, Boston, 1969, pp. 283–284.
31. Guilford, W. H., "Left Handedness: Its Effect Upon the Quality and Speed of Writing of Pupils in the Fifth and Sixth Grades," unpublished master's thesis, College of Puget Sound, Tacoma, Washington, 1936.
32. Harris, Theodore L.; Herrick, Virgil E., "Children's Perception of the Handwriting Task," in Herrick, V. E. (ed.), *New Horizons for Research in Handwriting*, University of Wisconsin Press, Madison, 1963, pp. 159–179.
33. Harrison, Enid M., "The Brain Damaged Child and Writing Problems," *Academic Therapy*, 4:13–21, 1968.
34. Herrick, Virgil E., *Comparison of Practices in Handwriting Advocated by Nineteen Commercial Systems of Handwriting Instruction*, Committee on Research in Basic Skills, University of Wisconsin, Madison, 1960.
35. Herrick, Virgil E., "Handwriting and Children's Writing," *Elementary English*, 37:248–258, April 1960.
36. Herrick, Virgil E., "Handwriting Tools for Children," *National Education Association*, 50:49–50, 1961.
37. Herrick, Virgil E.; Erlebacher, Adrienne, "*The Evaluation of Legibility in Handwriting*," in Herrick, V. E. (ed.), *New Horizons for Research in Handwriting*. University of Wisconsin Press, Madison, 1963, pp. 207–236.
38. Herrick, Virgil E.; Harris, Theodore, L.; Rarick, G. Lawrence, *Perception of Symbols in Skill Learning by Mentally Retarded, Gifted, and Normal Children*, Contracts SAE 6436 and 7135, Department of Health, Education, and Welfare, U. S. Office of Education, completed by Committee for Research in Basic Skills, School of Education, University of Wisconsin, Madison, 1961.
39. Herrick, Virgil E.; Okada, Nora, "The Present Scene: Practices in the Teaching of Handwriting in the United States," in Herrick, Virgil E. (ed.), *New Horizons for Research in Handwriting*, University of Wisconsin Press, Madison, 1963, pp. 17–32.
40. Herrick, Virgil E.; Otto, Wayne, *Letter Form Models Advocated by Commercial Handwriting Systems*, School of Education, University of Wisconsin, Madison, 1961.
41. Herrick, Virgil E.; Otto, Wayne, "Pressure on Point and Barrel of a Writing Instrument," *Journal of Experimental Education*, 30:215–30, December 1961.
42. Hildreth, Gertrude, "Manuscript Writing After Sixty Years," *Elementary English*, 37:3–13, May 1960.
43. Horn, Thomas D., "Handwriting and Spelling," *Review of Educational Research*, 37:168–177, 1967.
44. King, Fred M., "Handwriting Practices in Our Schools Today," *Elementary English*, 38:483–486, 1961.
45. Kuennapas, T., "Visual Perception of Capital Letters; Multidimensional Ratio Scaling and Multidimensional Similarity," *Scandinavian Journal of Psychology*, 7: 189–196, 1966.
46. Kuennapas, T., "Visual Memory of Capital Letters; Multidimensional Ratio Scaling and Multidimensional

Similarity," *Perceptual and Motor Skills*, 25:345–50, 1967.

47. Kuennapas, Teodor; Janson, Anne-Jeannette, "Multidimensional Similarity of Letters," *Perceptual and Motor Skills*, 28: 3–12, 1969.

48. Lewis, Edward R., "An Analysis of Children's Manuscript Handwriting," *Dissertation Abstracts*, 25:(no. 3) 1786–87, 1964.

49. Lewis, Edward R.; Lewis, Hilda P., "Which Manuscript Letters are Hard for First Graders?" *Elementary English*, 41:855–58, December 1964.

50. Lewis, Edward R.; Lewis, Hilda P., "An Analysis of Errors in the Formation of Manuscript Letters by First-Grade Children," in Otto, W.; Koenke, K. (eds.), *Remedial Teaching: Research and Comment*, Houghton Mifflin, Boston, 1969, pp. 273–282.

51. Lindell, Ebbe, *The Swedish Handwriting Method*, Ejnar Munksgaard, Copenhagen, Denmark, 1964.

52. Linn, Shirley, "Remedial Approaches to Handwriting Dysfunction," *Academic Therapy*, 4:43–46, 1968.

53. Love, Harold D., "Comparison of Quality, Speed, and Use of Handwriting Among Special and Regular Classroom Children," *Journal of Educational Research*, 58:475–477, 1965.

54. Lurçat, Liliane, "Etude de la Succession des Mots et du Contrôle de Leur Alignement dans une Épreuve de Dictée" (Study of the Succession of Words and of Control of their Alignment in a Dictation Test), *Journal de Psychologie Normale et Pathologie*, 65: 209–231, 1968.

55. McElravy, Anna, "Handwriting and the Slow Learner," *Elementary English*, 41:865–868, 1964.

56. Meis, Rudolf, *Schreibleistungen von Schulanfängern und das Problem der Anfangsschrift* (Performance in Writing of Children Entering School and the Problem of the Beginner's Script), C. J. Hogrefe, Göttingen, Germany, 1963.

57. Newland, T. Ernest, "An Analytical Study of the Development of Illegibilities in Handwriting from the Lower Grades to Adulthood," *Journal of Educational Research*, 26:249–258, December 1932.

58. Noble, J. Kendrick, "Handwriting Programs in Today's Schools," *Elementary English*, 40:513–17, May 1963.

59. Otto, Wayne; Andersen, Dan W., "Handwriting," in Ebel, R. L. (ed.), *Encyclopedia of Educational Research*, Fourth Edition, Macmillan, New York, 1969, pp. 570–579.

60. Otto, Wayne; Askov, Eunice; Cooper, Carin, "Legibility Ratings for Handwriting Samples: A Pragmatic Approach," in Otto, W.; Koenke, K. (eds.), *Remedial Teaching: Research and Comment*, Houghton Mifflin, Boston, 1969, pp. 272–273.

61. Otto, Wayne; Rarick, G. Lawrence, "Effect of Time of Transition from Manuscript to Cursive Writing upon Subsequent Performance in Handwriting, Spelling, and Reading," *Journal of Educational Research*, 62: 211–216, 1969.

62. Otto, Wayne; Rarick, G. Lawrence; Armstrong, Jenny; Koepke, Mariyln, "Evaluation of a Modified Grip in Handwriting," *Perceptual and Motor Skills*, 22:310, 1966.

63. Parres, Richard M., "The Body Midline as a Psychoeducational Variable Affecting the Handwriting Habits of Mentally Retarded Children," *Dissertation Abstracts*, 29:(no. 10)3468–3469-A, 1969.

64. Perron, Roger; Mignard, Edith, "Développement Graphique, Motricité et Intelligence: Contribution au Problème par l'Étude des Débiles Mentaux" (Graphic Development, Moticity, and Intelligence: Contribution to the Problem by a Study of Mental Retardates), *Enface*, No. 5, 1965, pp. 532–552.

65. Petty, Walter T., "Handwriting and Spelling: Their Current Status in the Language Arts Curriculum," in Horn, Thomas A. (ed.), *Research on Handwriting and Spelling*, National Council of Teachers of English, Champaign, Illinois, 1966, pp. 1–8.

66. Plattor, Emma E., "The Development of a Programmed Course of Study in Manuscript and Cursive Handwriting for Use in Teacher Education," *Dissertation Abstracts*, 27:(no. 1)131-A, 1966.

67. Pressey, Sidney L.; Pressey, Luella C., "Analysis of 3,000 Illegibilities in the Handwriting of Children and Adults," *Education Research Bulletin*, 6:270–273, 1927.

68. Quant, Leslie, "Factors Affecting the Legibility of Handwriting," *Journal of Experimental Education*, 14:297–316, June 1946.

69. Rarick, G. Lawrence; Harris, Theodore L., "Physiological and Motor Correlates of Handwriting Legibility," in Herrick, V. E. (ed.), *New Horizons for Research in Handwriting*, University of Wisconsin Press, Madison, 1963, pp. 55–94.

70. Reed, G. F.; Smith, A. C., "A Further Experimental Investigation of the Relative Speeds of Left and Right Handed Writers," *Journal of Genetic Psychology*, 100: 275–288, 1962.

71. Rondinella, Oreste R., "An Evaluation of Subjectivity of Elementary School Teachers in Grading Handwriting," *Elementary English*, 40:531–532, 1963.

72. Schell, Leo M.; Burns, Paul C., "Retention and Changes by College Students of Certain Upper-case Cursive Letter Forms," *Elementary English*, 40:513–17, May 1963.

73. Seifert, Eloise P., "Personal Styles of Handwriting in Grades 6, 7, 8, and 9," *Dissertation Abstracts*, 20:(no. 9)3581–82, 1960.

74. Shane, Harold G.; Mulry, June G., *Improving Language Arts Instruction Through Research*, Association for Supervision and Curriculum Development, National Education Association, Washington, D. C., 1963, pp. 45–60.

75. Smith, K. U.; Smith, W. M., *Perception and Motion*, Saunders, Philadelphia, 1962.

76. Sokolova, E. N., "Development of Writing Habits in Primary School Children," *Voprosy Psikhologii*, No. 5, pp. 140–146, 1963.

77. Tagatz, Glenn E.; Otto, Wayne; Klausmeier, Herbert J.; Goodwin, William L.; Cook, Doris M., "Effect of Three Methods of Instruction Upon the Handwriting Performance of Third and Fourth Graders," in Otto, W.; Koenke, K. (eds.), *Remedial Teaching: Research and Comment*, Houghton Mifflin, Boston, 1969, pp. 234–239.

78. Templin, Elaine M., "The Legibility of Adult Manuscript, Cursive, or Manuscript-Cursive Handwriting, Styles," in Herrick, Virgil E. (ed.), *New Horizons for Research in Handwriting*, University of Wisconsin Press, Madison, 1963, pp. 185–200.

79. *Times Education Supplement*, "New Handwriting; Russian Style," 2641:1391, 1965.

80. Townsend, Edward A., "A Study of Copying Ability in Children," *Genetic Psychology Monographs*, 43:3–51, 1951.

81. Weinert, F.; Simons, H; Essig, W., *Theorie und Praxis der Schulpsychologie, Band V: Schreiblehrmethode und Schreibentwicklung* (Theory and Practice of the School Psychology, Volume 5, The Teaching of Handwriting and its Development), Verlag Julius Beltz, Weinheim, Germany, 1966.

82. "Writing Device Project by Industrial Design Students at Syracuse University," *Industrial Design*, 11:66–67, 1964.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

The EPPS Need Structure of In-Service Elementary School Teachers

E. LARRY EBERLEIN, University of Alberta

ABSTRACT

The Edwards Personal Preference Schedule (EPPS) was administered to 293 female Michigan elementary school teachers. This study replicates a portion of Jackson and Guba's study (5) and considers factors of maturation (age and experience) and the effects of the EPPS forced-choice format. N-Deference showed an increase with age; n-Exhibition and n-Heterosexuality showed a decrease. Needs for Order, Endurance, Abatement were greater in older teachers. When compared with adult women, the rank-order correlation was .87 in spite of numerical differences on eleven of fifteen needs. The effect of rank-ordering on interpretation of differences in group comparisons is considered.

THE EPPS, a personality test widely used for research, was designed to compare the relative strength of fifteen manifest needs. Jackson and Guba (5) utilized this test to study public school teachers with varying amounts of teaching experience in suburban areas of Illinois, Wisconsin, and Indiana, with the results reported in terms of college-age norms. The present study is an effort to replicate a portion of the Jackson and Guba study, and analyze the data using the newer adult women's norms reported by Edwards (2).

Maturation is important while considering the effects of age and experience in relation to teacher need structure. In comparing seventy-nine female teachers having less than 4 years experience with ninety-three veteran teachers, who had more than 10 years experience, Jackson and Guba found the veteran groups significantly higher on needs for Deference, Order, Endurance, and Aggression, but lower on Heterosexuality and Exhibition. They suggested some differences might have been due to age rather than vocation.

Other studies have since indicated that greater weight needs to be given maturation. Differences are reported, for example, by Klett (6) and Edwards (2) among high school students, college students, and the general population. Koponen (7) apparently found demographic differences in his sample from which EPPS adult norms were derived. Plapp, Psathas, and Caputo (8) concluded that nursing training possibly impeded certain age-related need changes. Using the "recurrent institutional cycle design" in an effort to eliminate sex and maturation, the only significant differences related to the nursing program were on needs for Deference and Heterosexuality. Spangler and Thomas (9) found that scores of Affiliation, Deference, and Succorance tended to increase with age. Gauron (3) gave the EPPS to patients upon their admission to a psychiatric hospital and found that eight of the needs showed a relationship to age similar to that reported in the Minnesota Multiphasic Personality Inventory (MMPI) literature.

In the present study the EPPS was administered to 315 teachers in twenty-four Michigan elementary schools. The vast majority of the sample was currently married; fourteen were separated or divorced; twenty-one were widows, and forty-two were still single. In terms of vocational aspirations, 41 percent planned to teach until retirement; 22 percent planned to teach for the foreseeable future; 13 percent saw themselves teaching less than 5 years now, but coming back to education later. The remaining 24 percent had plans for non-classroom participation in education or viewed teaching as a temporary job. All principals and male teachers were eliminated from consideration, as they were more likely to have an administrative profile. For 293 female teachers who completed the EPPS, correlations were computed between each of the fifteen manifest needs and the teachers' age and experience (see Table 1). The sample was sequentially divided into (a) five 10-year age groupings, (b) three groups based upon experience in education, and (c) three groups based on experience in the present school building. Rank order comparisons were made using the Klett and Edwards' data along with the present 10-year age groupings. Group average EPPS scores were compared with those obtained by Klett, Edwards, and Jackson and Guba, and rankings were correlated (see Table 2).

Table 1.—EPPS Numerical Values for 293 Female Elementary School Teachers and Correlations Between EPPS Needs and Age, Experience in Education, and Experience in Building

EPPS Need	Value	Correlations Age	Correlations Exp. Educ.	Correlations Exp. Bldg.
Achievement	13.38	-.06	-.09	-.04
Deference	14.71	.42	.44	.28
Order	13.19*	.21	.23	.12
Exhibition	12.98*	-.28	-.32	-.22
Autonomy	12.87*#	-.03	-.05	-.01
Affiliation	17.19	.12	.13	.09
Intraception	16.95*	-.05	-.07	-.00
Succorance	11.93*	-.08	-.02	-.08
Dominance	11.25*	-.07	-.14	-.07
Abasement	15.06*	.22	.28	.15
Nurturance	15.94*	.12	.12	.14
Change	17.52*	-.16	-.16	-.09
Endurance	15.28*	.32	.31	.19
Heterosexuality	11.18*	-.39	-.38	-.27
Aggression	10.52	-.10	-.11	-.07
Age		1.00	.84	.55
Experience Educ.			1.00	.65
Experience Bldg.				1.00

*p < .01 when compared with Edwards Adult Women Norms.
#p < .01 when compared with Jackson & Guba (1957) data.

Results

When compared with women in general (Edwards' data), women elementary school teachers in the present sample were found to be significantly higher on need for Exhibition, Autonomy, Intraception, Dominance, Change, and Heterosexuality. They were significantly lower on Order, Succorance, Abasement, Nurturance, and Endurance.

The present study replicated the Jackson and Guba data except for n-Autonomy. Jackson and Guba found their sample of teachers to have less need for autonomy than the adult norms suggested for women in general. The present study showed the reverse. This is the only score significantly different from the Jackson and Guba data. Since there appears to be less need manifested for teachers to depend on others (or to be depended upon by others), it is reasonable to expect that they would also show a desire to be more autonomous, as indicated in the present study.

Subject to the limitations suggested below, significant differences might be interpreted as follows. When compared with women in general, women who teach in an elementary school manifest a desire to be the center of attention and their own bosses. They are more likely to need to analyze their own and other's motives and behavior, and to do new and different things. A greater need to dominate, persuade, and influence others to do as they desire is present, as is a desire to engage in activities with men. On the other hand, female elementary school teachers manifest less need to keep things in order, be well organized, and plan in advance. They indicate less desire to have others provide help and encouragement, and also have less need to help their friends when in trouble (and thus to be more independent and autonomous themselves). There is likely less need to feel guilty, accept blame when things go wrong, or to stick with a job until it is finished. This is in contrast with the stereotype of teachers as "sexually impotent, obsequious, eternally patient, painstakingly demanding, and socially inept" (5:189).

Six needs showed a definite relationship to age and experience. Deference showed a consistent linear increase with maturation. Needs for Exhibition and Heterosexuality showed decreases with age and experience, although the latter tended to stabilize at about age 30, and drop sharply after age 50. At age 40, teachers tended to show sharp increases (and then stability) in needs for Order and Endurance. The same trend occurred about age 50 in the need for Abasement. Inspection of the rank ordering of needs by age tends to confirm these trends. Needs for Change, Nurturance, Intraception, and Affiliation tend to rank consistently high at any age. N-Aggression and Dominance consistently remain toward the lower end of the list. The strongest trends appeared in the need for Deference, which showed a linear increase in both comparative rank and average value, and in the needs for Exhibition and Heterosexuality which showed decreases.

Discussion

This study supports other evidence that individual needs, as measured by the EPPS, showed systematic patterns of change related to maturation. The study was cross-sectional and not longitudinal. Other factors such as experimental selection bias, mortality, marital status, health, and the degree to which individual needs were being met (a problem when manifest needs are being considered) are involved along with age and experience in producing these results.

Interpretation of EPPS scores is also hindered by the ipsative, forced-choice format of the instru-

Table 2.—Rank Order Correlations on Female EPPS Scores

	Eberlein Teachers	Jackson & Guba (1957)	Edwards (1959) Adults	Edwards (1959) College	Klett (1957) High School
Eberlein	1.00	.95	.87	.65	.58
Jackson & Guba		1.00	.82	.68	.59
Edwards: Adult			1.00	.44	.45
Edwards: College				1.00	.89
Klett					1.00

ment. The EPPS rank-orders individual needs on an ordinal scale based on each individual's comparative judgment. Radcliffe and Stricker (1: 190–207) question the implications of ordinal ipsative scores. Radcliffe, for example, notes that rank orders of needs may change considerably when converted and expressed in normative form. Heilbrun (4) claimed on the basis of moderate correlations that ipsative measures could be used for normative predictions. In large sample sizes, however, small differences tend to be statistically significant, even though of questionable practical significance. Data from the present study show female elementary school teachers differing from adult women in general on eleven of the fifteen needs. It replicated Jackson and Guba's findings on all but the need for Autonomy. Even with replicated results, it is difficult when using forced-choice data to adequately interpret rank-ordered scores.

When needs are rank-ordered and interpreted in terms of age and experience, one finds n-Autonomy and n-Dominance in the same relative position in need hierarchies produced from both the present data and Edwards' adult women norms. Needs for Endurance and Heterosexuality are only one rank apart. The greatest discrepancies are in needs for Interception and Change, both being four ranks apart. The rank order correlation between these two lists is .87, surely a substantial agreement, in spite of eleven statistically significant differences between average values.

The largest difference between means was in the need for Heterosexuality. The present sample included eighty-eight married and twenty-six single teachers in the 20–29 age category. Since the need for Heterosexuality dropped sharply at age 30, the elimination of 114 younger teachers reduced the average manifest need for Heterosexuality from 11.81 to 9.51. This result is more in line with the Edwards' Adult Female Norms, and provides a good example of how average data are affected by a large number in a single age category.

The need for Nurturance is ranked highest in Edwards' female norms, a quality becoming a woman who is the head of the household. The fact that elementary school teachers' average score is substantially less suggests a relationship with work experience and should be replicated in most groups of working women. In terms of rank order, however, the need for Nurturance is still consistently high on the teachers' need list.

Conclusion

The interpretation and generalization of EPPS average scores becomes more difficult as more factors are considered. In this cross-sectional study of female teachers, manifested needs change with age and experience. A longitudinal study may have shown different results. Where groups differ on average values for a specific need, strong similarities may yet exist when needs are compared on a rank order basis. At the same time, rank changes may not be important when the magnitude of adjacent needs are similar. Consideration must be given to both absolute and relative values when interpreting the EPPS forced-choice format. It is suggested that age and experience factors be controlled, and EPPS data compared on both average value and rank order basis, or some combination of both.

REFERENCES

1. Buros, O. K., *The Sixth Mental Measurements Yearbook*, Gryphon Press, Highland Park, New Jersey, 1965.
2. Edwards, A. L., *Edwards Personal Preference Schedule; Manual. (Rev. Ed.)*, Psychological Corporation, New York, 1959.
3. Gauron, E. F., "Changes in Edwards Personal Preference Schedule Needs with Age and Psychiatric Status," *Journal of Clinical Psychology*, 21:194–196, 1965.
4. Heilbrun, A. B., Jr., "Evidence Regarding the Equivalence of Ipsative and Normative Personality Scales," *Journal of Consulting Psychology*, 27:152–156, 1963.
5. Jackson, P. W.; Guba, E. G., "The Need Structure of In-Service Teachers: An Occupational Analysis," *School Review*, 65:176–192, 1957.
6. Klett, C. J., "The Performance of High School Students on the Edwards Personal Preference Schedule," *Journal of Consulting Psychology*, 21:68–72, 1957.
7. Koponen, A., "The Influence of Demographic Factors on Responses to the Edwards Personal Preference Schedule," unpublished doctoral dissertation, Columbia University, 1957.
8. Plapp, J. M.; Psathas, G.; Caputo, D. V., "Effects of Nursing Training on Personality," *Proceedings of the 74th Annual Convention of the American Psychological Association*, 1966, pp. 287–288.
9. Spangler, D. P.; Thomas, C. W., "The Effect of Age, Sex, and Physical Disability Upon Manifest Needs," *Journal of Counseling Psychology*, 9:313–319, 1962.

BOOK REVIEWS

ROBERT E. CLASEN
Book Review Editor

i.t.a.: An Independent Evaluation

Warburton, F. W.; Southgate, Vera. John Murray and W. and R. Chambers, London, 1969, 321 pp. $12.00.

THE EXTENT of use of i.t.a. (initial teaching alphabet), of primary teachers' and other interested persons' reactions to the impact of i.t.a., and of seventeen independent investigations of i.t.a. in Great Britain and the United States is encompassed in *i.t.a.: An Independent Evaluation*. This is the book-length report of a study carried out for the Schools Council in Great Britain. Although the bulk of the data were gathered in Great Britain, where i.t.a. has been quite extensively adopted

(Continued on page 126)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

A Model for Teaching Composition

LESTER S. GOLUB, University of Wisconsin

ABSTRACT

A nontraditional model for teaching composition has extensive teaching and research value. The purpose of this research was to test the effect of oral discourse in teaching written discourse. The control group went from *stimulus* to *problem* to *written discussion;* the experimental group went from *stimulus* to *problem* to controlled *oral discussion* to *written discussion* on the model. Teachers evaluated unidentified papers of both pretest and posttest of control and experimental groups. With the use of the chi-square statistic, it was shown that the use of controlled oral discourse produced more grade increases and fewer decreases in written discourse in the experimental group than in the control group at the .01 level of significance.

EDUCATORS and researchers realize that, at the end of high school or the first year of college, students will not write with the same degree of competence. Yet there is no reason why an individual student cannot develop a competent oral and written style which is compatible with his own environment, age, sex, and socioeconomic background. Teachers of composition need not assume that writing cannot be taught simply because very few Orwells or Emersons emerge from our English classrooms. Instead, teachers should strive for students who can express, either orally or in writing, their unique ideas and emotions in an honest, understandable fashion. Goals for teaching composition can be stated as follows:

(1) The student should develop a positive attitude and motivation for expressing his thoughts and perceptions in oral and written language.

(2) He should produce writing which is clear, direct, economical, and sincere.

(3) He should be able to write clearly and logically on a variety of concepts and emotions relevant to him in a variety of ways.

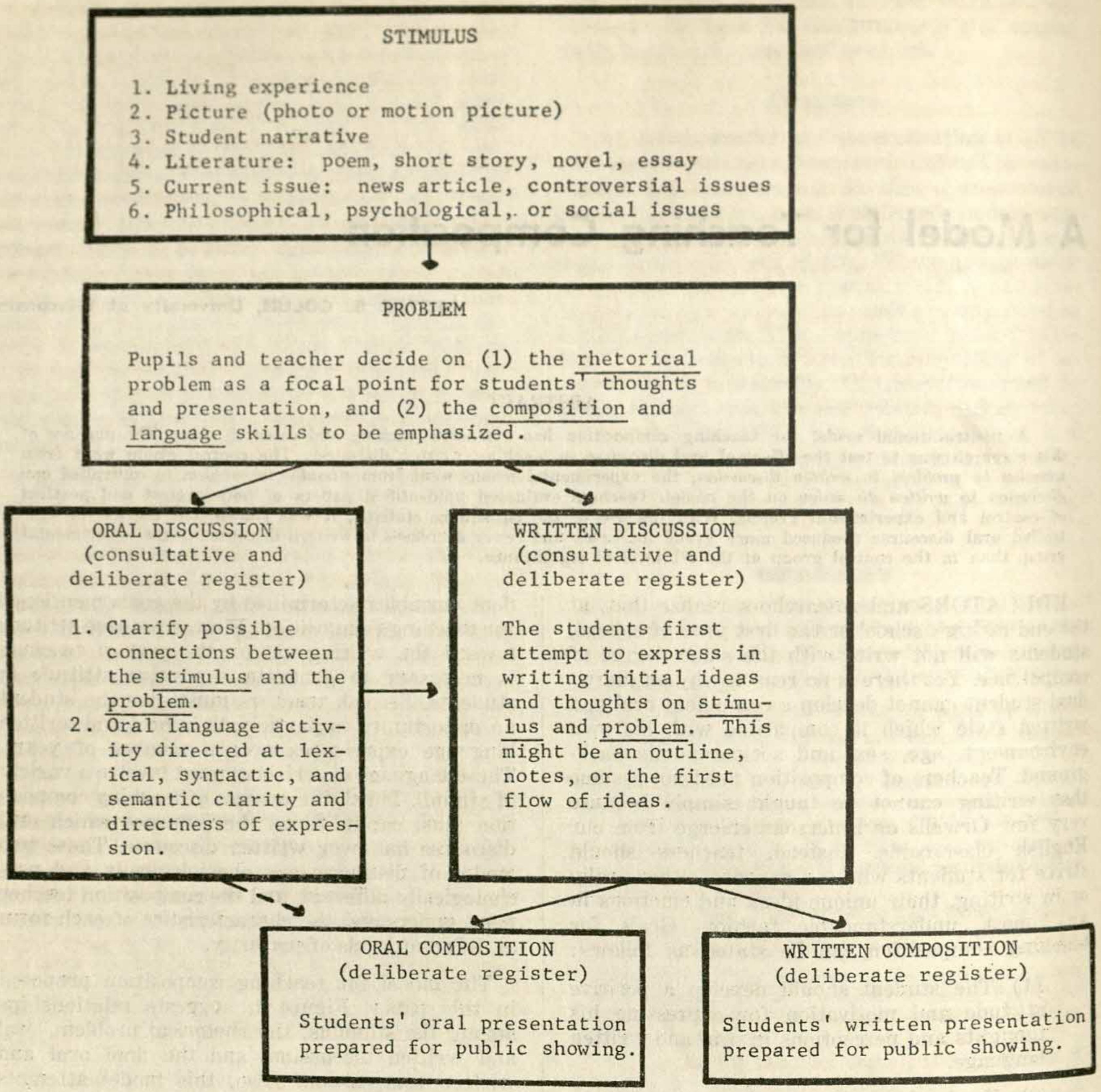
The traditional composition model in the United States has been the write-correct-rewrite model which lacks in teaching flexibility and creativity. This deficiency prevents students and teachers from achieving the established goals. The results of the Anglo–American Conference at Dartmouth College in 1966 imply that a significantly different model for teaching composition is needed.

Such a model for teaching composition must consider the classroom factors and teacher–student variables determined by the goals mentioned for teaching composition. First, a positive attitude toward the writing act on the part of teachers is necessary to generate a positive attitude in students. Second, teachers must give the student an opportunity to participate in oral and written language experiences over a number of years. These language experiences must involve a variety of stimuli. Third, the process of teaching composition must capitalize on the influence which oral discourse has over written discourse. These two modes of discourse are physiologically and psychologically different, and the composition teacher must understand the characteristics of each form at different levels of maturity.

The model for teaching composition proposed in this paper, Figure 1, suggests relationships among the stimulus, the rhetorical problem, oral and written discussion, and the final oral and written composition. Also, this model attempts to utilize linguistic, rhetoric, and composition techniques and knowledge available to teachers and researchers. The diagram is a schematic presentation of the kinds of teacher–student composition activities which can take place at any grade level. In the following discussion of the model, the solid arrows indicate an uninterrupted flow of activities; the broken arrows indicate possible branchings.

The *stimulus* for the written or oral expression is the most urgent need for the student. The stimulus must be relevant to both student and teacher. It might come from a living experience, a photograph or movie, a student narration, a current issue, a work of literature, or some pertinent problem. At an elementary or junior

FIGURE 1. A model for teaching composition


STIMULUS
1. Living experience
2. Picture (photo or motion picture)
3. Student narrative
4. Literature: poem, short story, novel, essay
5. Current issue: news article, controversial issues
6. Philosophical, psychological, or social issues
PROBLEM
Pupils and teacher decide on (1) the rhetorical problem as a focal point for students' thoughts and presentation, and (2) the composition and language skills to be emphasized.
ORAL DISCUSSION (consultative and deliberate register)
1. Clarify possible connections between the stimulus and the problem.
2. Oral language activity directed at lexical, syntactic, and semantic clarity and directness of expression.
WRITTEN DISCUSSION (consultative and deliberate register)
The students first attempt to express in writing initial ideas and thoughts on stimulus and problem. This might be an outline, notes, or the first flow of ideas.
ORAL COMPOSITION (deliberate register)
Students' oral presentation prepared for public showing.
WRITTEN COMPOSITION (deliberate register)
Students' written presentation prepared for public showing.

high level the stimulus might serve mainly to activate the flow of discourse. Much of the response to this kind of stimulus should be oral and dramatic as well as written. In high school or college the stimulus may become more sophisticated and serve as the vehicle for the presentation and organization of concepts, as well as the vehicle for the development of new, expanded concepts and values.

The following example of student writing illustrates stereotyped, artificial prose written to satisfy the student's belief of teacher expectations:

> *I think the countryside is truly magnificent. It allows one to escape from the everyday bustle of the city life and to enter into a beautiful garden of tranquility.*
>
> Grade 12

The sample of student prose below represents sincere, serious expression written in a clear, logical, direct economical style:

> *The loneliest place in the world is a recovery room after surgery. But this kind of loneliness is different than the kind you know. Even with all the people there, you're lonely because the people have to be there. They don't necessarily want to be there. Maybe there isn't anyone in the room, but then a nurse comes into a room and sticks a needle into your arm, even though you tell her you aren't having pain and the tears you were crying weren't tears of pain. You really can't tell her why you're crying, because she's one of the people who has to be there. The only way you can get over this feeling is being with someone who wants to be there, who wants to try to cheer you. Not your parents, but any kind of close friend.*
>
> Grade 7

A *problem* must accompany each stimulus in order to generate creative thinking. The problem can relate to rhetoric, language, and composition skills. For example, a teacher produced from a brown paper bag stimuli in the form of an old high-top boot, a carved wooden figure of an aging man, a 1920 alarm clock, and a stuffed frog and suggested as the problem, "What do these objects remind you of from some past experience?" A student who had shown no previous writing ability produced the following poem:

A Frog's Too Real

When I was eight and thought that I
Could do what any boy could do,
Long summer days were really made
For fishing in the green canoe.

* * *

But then the boys decided that
The time had come, since I was eight,
When I should learn to cast for bass.
And frogs were what we'd use for bait.

* * *

Then I took a frog from the minnow pail
But when I brought the hook up near
I couldn't bear to thrust it in.
His throat pulsed hard, it seemed,
with fear;
His popping eyes were frightened too.
I somehow knew what pain he'd feel.
I couldn't stand to have him die
Skewered on that hook of steel.
So I turned him loose, and didn't care
That I'd failed the test one should pass
at eight.
I'll fish with worms and minnows still
But a frog's too real to use for bait.

Grade 8

The more mature students can be presented problems such as audience appeal, argument, persuasion, approaches to logic, and paragraph development. Language problems include word production, vocabulary development, use of figurative language, use of coordination and subordination, creative use of movable linguistic elements, and expansions of the verb string. During the presentation and discussion of the problem, writing of peers as well as professional writers can be discussed, but care must be taken not to overwhelm the students' creative imaginations by offering examples of professional writing too difficult for students to emulate.

Oral discussion is not found in the traditional composition model. In the model presented in this paper, oral discussion is of two kinds. One concerns the correspondence between the stimulus and the rhetorical or language problem. Questions will be asked such as "How does it feel to eat a carrot?" or "What would it be like to be a poor Mississippi Delta Negro? Try to describe your condition in two letters, one to James Baldwin, the other to Senator Eastland." The second kind of discussion concerns some deliberate language study and direct experience with certain language skills. For example, the discussion of the problem, "How does it feel to eat a carrot?" calls for vocabulary pertaining to the senses, sight, smell, size, taste, touch, feel. It also needs words that get to the feeling component of eating a cool carrot on a warm day. The topic dealing with the human condition of a Mississippi Delta Negro might include some notions of dialect, vocabulary pertinent to the culture of poverty, tone of expression, and rhetorical effect of subordination and coordination.

The following poem illustrates clear expression of intimate feeling following a class discussion dealing with the problem of the inadequacy of language in some human situations.

Her

We sat apart from each other
Our feelings overwhelming,
We did not speak,
We sat in complete silence.

I turned to speak,
She lifted her head,
But I, like a coward I am,
Turned away, and hope fell to the depths.

She talked about the weather,
Who cared about the weather?
But we talked,
The barrier was broken,
We walked off hand in hand.

Grade 10

Written discussion in the model provides for the students' initial attempt to write after the stimulus, the problem, and the oral discussion have been presented. This written discussion might be the initial flow of ideas, the first draft, an outline, or simply some random notes or ideas on the subject. At this phase of the model, the teacher and other classmates become involved in a critique of the initial written response by asking: How can the ideas be better organized? How can the mechanical problems (spelling, punctuation) be altered? What language changes or revisions might be helpful? Is the topic important enough for the student to respond honestly to it?

From the initial, informal writing and discussion, the student can go to the more formal, polished *oral or written composition*. Not all stimuli should be developed into *written compositions*. Some stimuli should result in *oral composition* in the deliberate register. The student should collect his work during the school year and eventually save the best for a final evaluation. Formal teacher evaluation of the compositions need not be made on the practice activities but rather on the final "show" of the products which the student, his teacher, and his peers consider his best performances. The evaluative criteria should then be the quality of the student's performance on principles emphasized during the period of instruction.

Traditionally, the teacher evaluates students' compositions. In the classroom this model is used, the teacher aids and directs students as an adult specialist and tries at the same time to improve his own composition as well as that of his students. Also, with this type of composition teacher, each student is involved in a responsible way with the performance of his classmates. This proposed model also forces teachers to look at language study as an oral activity as well as a written activity.

There are uses for the model in composition research. For example, a recent study has shown that the *Oral Discussion* between the *Problem* and the *Written Discussion* does make a difference in students' composition performance even though the broken arrow on the model indicates that this phase might be bypassed.

One hundred and twelve ninth-grade students were randomly divided into a control group (N = 56) and an experimental group (N = 56). Both groups were administered the same *Stimulus* and the same *Problem* treatment. At this point, the control group went directly from *Problem to Written Discussion*, whereas the experimental group went from *Problem* to *Oral Discussion* to *Written Discussion*. The experimental group had oral language activities which were designed to help students achieve the six rhetorical principles mentioned in the *Problem* portion of the treatment. Students verbalized possible interesting titles; they compared and contrasted similarities and differences of objects very much alike or different; they distinguished unique characteristics and tried to incorporate these unique characteristics into specific images by using figures of speech and content-specific vocabulary; they were given the technique for forming and using metaphor and simile; they were given opportunities to relocate nouns, verbs, adjectives, and adverbs; and they were given oral activities involving coordination and subordination procedures.

Three teachers who did not know which themes were from which treatments, evaluated, on a 7-point scale, pre-writing and post-writing of control and experimental groups. Table 1 shows that the increase in change of writing behavior of the experimental group was significantly greater than the increase in change of writing behavior of the control group at the .01 level of significance.

Table 1.—Observed and Expected Changes in Writing Behavior from Pre- to Post-writing Test

Group	Observed Increases	Observed Decreases	Observed No Change
Control (56 S)	21	12	23
Exper. (56)	42	4	11
Group	**Expected Increases**	**Expected Decreases**	**Expected No Change**
Control (56)	31.5	8	17
Exper. (56)	31.5	8	17

(X^2 = 14.6 $p < .001$ for 2 df)

Further research is needed to verify and implement the model for teaching composition. Teachers and psychologists must cooperatively determine the kinds of stimuli that can be used for specific purposes at different age and grade levels and with different ability and temperament groups. Teachers, psychologists, linguists, and rhetoricians should exchange ideas in determining the kinds of problems to propose with certain stimuli at grade levels and the kind of oral language experience that should accompany these stimuli and problems. Finally, teachers, professional writers, and lay readers should determine levels of competence at various grade, ability, and temperament levels. Certainly a highly literate, technical society demands oral and written expression from its citizens which is clear, direct, and logical and which creatively deals with a variety of concepts, values, and attitudes.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Effects of Training on the Instructional Behaviors of Student Tutors[1]

FRED C. NIEDERMEYER
Southwest Regional Laboratory for Educational Research and Development, Inglewood, California

ABSTRACT

The study assessed the instructional behaviors of trained student tutors and compared these behaviors with those of untrained tutors. Behaviors of six trained and six untrained fifth-grade tutors were recorded on the Tutor Observation Scale as the tutors worked with kindergarten pupils using programmed materials related to a reading program. Trained tutors scored significantly higher than untrained tutors on (1) engaging the pupil in friendly conversation, (2) confirming, (3) praising, (4) giving the correct answer when the pupil is incorrect, and (5) eliciting the correct response before going on. The behaviors of trained and untrained tutors differ with respect to basic instructional principles. If it is desired that tutors behave according to these principles, then they should be trained.

CURRENTLY there is considerable interest in the use of student tutors and other nonprofessionals in instruction. To what extent do these persons need to be trained? Do untrained tutors display the same instructional behaviors as those trained in specific tutorial skills?

To investigate the above questions, a Tutor Observation Scale (Figure 1) was developed. The Scale was then used by two observers to record the interaction between twelve fifth-grade tutors (6 trained and 6 untrained) and twelve kindergarten children at two schools. The tutors used programmed materials related to a kindergarten reading program.

The observation scale was designed to assess instructional behaviors related to the objectives of a tutor training program. Briefly, these behaviors were as follows:

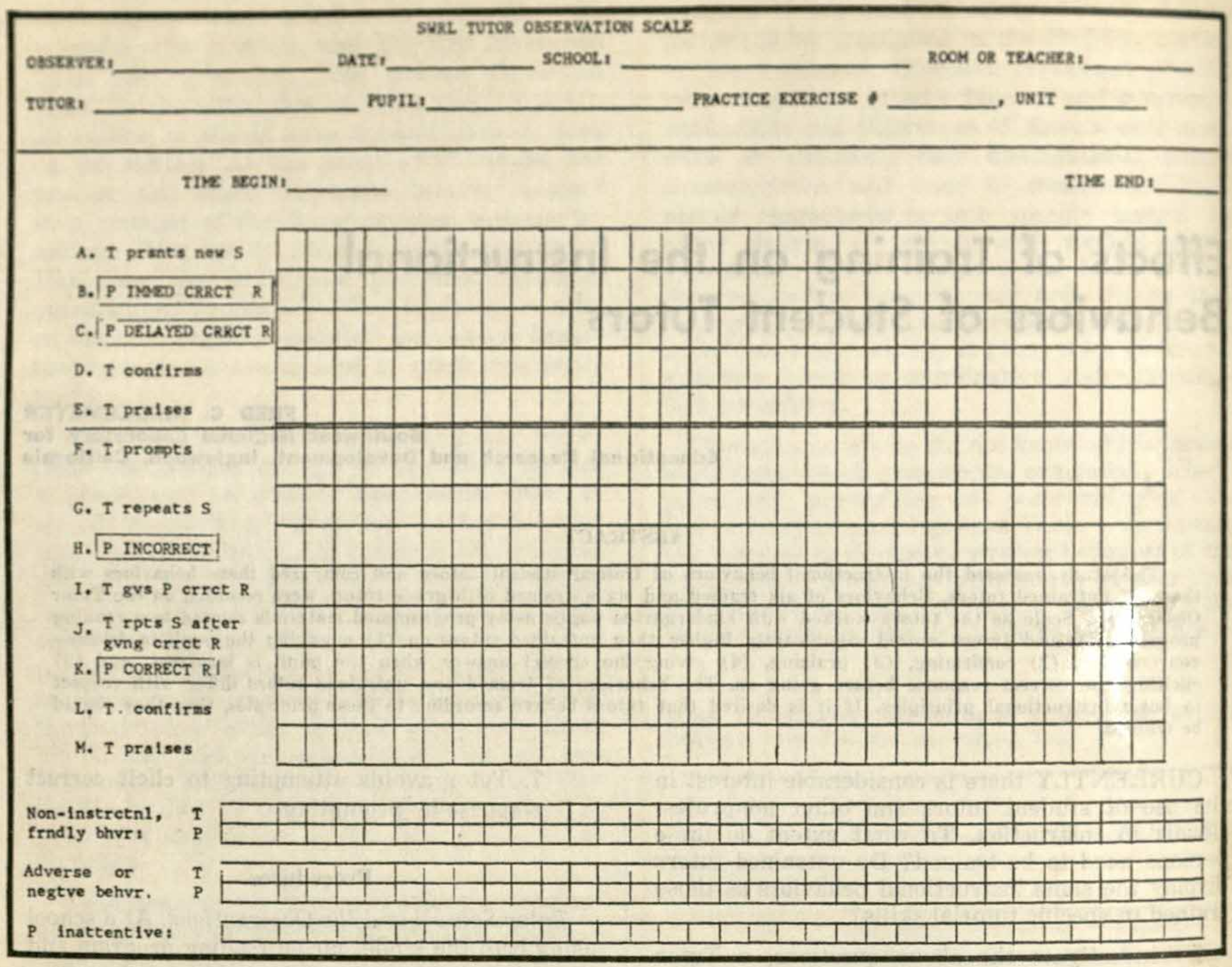
1. Tutor engages pupil in noninstructional, friendly conversation.
2. Tutor verbally confirms correct pupil responses.
3. Tutor praises the pupil.
4. Tutor tells or shows the pupil the correct response when the pupil is incorrect.
5. Tutor, after displaying Behavior 4, then elicits correct response from pupil before going on.
6. Tutor, following nonresponses to his initial question or direction, repeats it in different words.
7. Tutor avoids attempting to elicit correct response by prompting.

Procedures

Tutor Sample and the Observations. At a school using both the kindergarten reading program and the tutorial program, trained fifth-grade tutors were randomly selected for observation by two observers during a tutoring session. These students had been tutoring for about 4 weeks. They were told to conduct their session as usual, and that the observers were interested in "watching the kindergarten children to find out how they learn to read."

At a nearby school, which was using the reading program but not the tutorial program, fifth-grade students were randomly selected from pupils who, according to the teachers, would be acceptable as tutor trainees. As in the other school, the kindergarten pupils to be tutored were selected on the basis of their performance in the reading program. Prior to the sessions, each untrained tutor was shown the materials (called Practice Exercises) he was to use. He was told to read the script to the child and "help him learn the words and sounds in the Exercise." He was also given the opportunity to ask questions. As with the trained tutors, untrained tutors were told that the observers were primarily interested in the kindergarten pupils.

Figure 1. SWRL Tutor Observation Scale



SWRL TUTOR OBSERVATION SCALE

OBSERVER: ________ DATE: ________ SCHOOL: ________ ROOM OR TEACHER: ________

TUTOR: ________ PUPIL: ________ PRACTICE EXERCISE # ________, UNIT ________

TIME BEGIN: ________ TIME END: ________

A. T prsnts new S

B. P IMMED CRRCT R

C. P DELAYED CRRCT R

D. T confirms

E. T praises

F. T prompts

G. T repeats S

H. P INCORRECT

I. T gvs P crrct R

J. T rpts S after gvng crrct R

K. P CORRECT R

L. T. confirms

M. T praises

Non-instrctnl, frndly bhvr: T P

Adverse or negtve behvr. T P

P inattentive:

Two observers were present at each of the tutoring sessions, and each observer independently recorded the tutorial interaction on the Tutor Observation Scale.

Instructional Materials. Figure 2 is a typical page from one of the 5-page Practice Exercises used by tutors with the kindergarten pupils. For each of the four items on the page, the tutor presents the initial verbal stimulus to the child by reading the vertically printed script on the left. Thus these items, each one a clearly defined stimulus–response unit, formed the basis for developing an observation scale to assess tutor–student interaction.

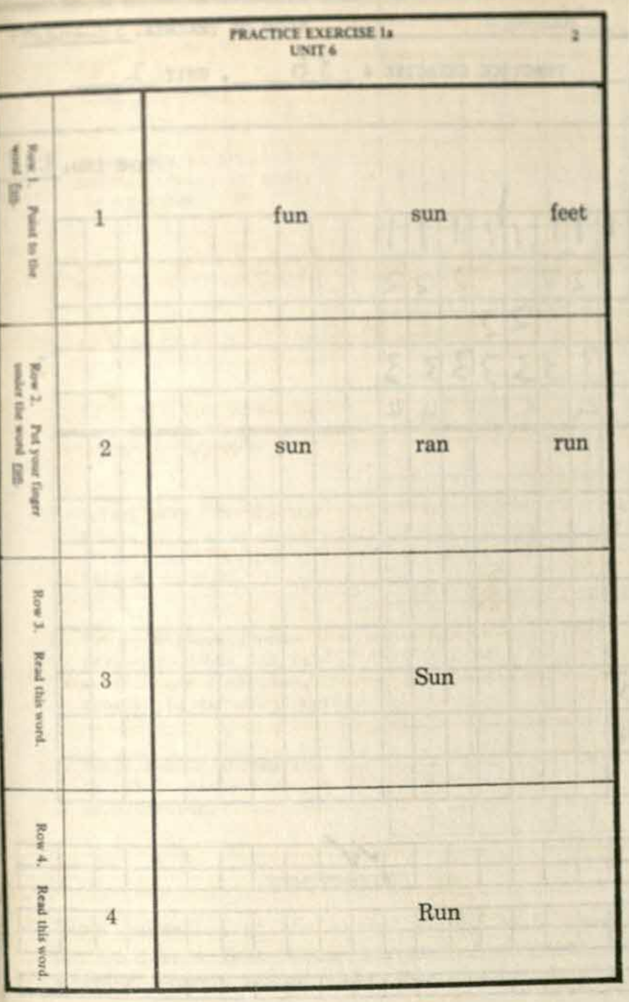
The Tutor Observation Scale. Figure 3 is a completed scale that was used when observing a trained tutor with a kindergarten child. On the left, lettered *A* through *M*, are various tutor (T) or pupil (P) behaviors (pupil behaviors are boxed and in capitals). Each column in the grid represents a stimulus–response (S–R) unit, i.e., an item in the Practice Exercise. The history of tutor–pupil behaviors on each S–R unit is recorded in a column on the grid as the observer sequentially numbers the listed behaviors when they occur. When the tutor and the pupil move on to another item in the Practice Exercise, the observer simply begins numbering again in the next column on the grid.

Note, for example, the first column on the left side of the grid. The numerals recorded in this column represent the sequence of tutor–pupil behaviors as they occurred on the first item in the Practice Exercise. The verbal interaction may have sounded like this:

> Tutor presents stimulus, "Read this word." (Observer writes "1.")
> Pupil responds correctly (after delay of 3 seconds), "Run." (Observer writes "2".)
> Tutor confirms, "That's right." (Observer writes "3," and moves to the next column.)

The variations in the numbering patterns from

Figure 2. Typical page from a practice exercise used by tutors with kindergarten pupils



PRACTICE EXERCISE 1a
UNIT 6
2

Row 1. Point to the word fun.	1	fun sun feet
Row 2. Put your finger under the word run.	2	sun ran run
Row 3. Read this word.	3	Sun
Row 4. Read this word.	4	Run

column to column represent differences in the tutor–pupil chain of behaviors on different items. The numbers in the third column, for example, indicate a more involved sequence than in the first column. This is because the pupil made an incorrect response on the third item in the Practice Exercise.

After the observation is completed, the tutor's score on various behaviors of interest can be calculated, e.g., the proportion of correct pupil responses verbally confirmed by tutor.

Results

The instructional behaviors of twelve tutors (6 trained, 6 untrained), as recorded with the scale, are presented in Table 1.

Table 1 shows the degree to which each of the six trained tutors and six untrained tutors displayed seven behaviors derived from the objectives of the training program. Totals and means (percentage where appropriate) were computed for each of the two groups of tutors. A rank order correlation of agreement (Spearman's *Rho*) between the two observers across the twelve tutors is listed for each behavior. All are significant beyond the .01 level. A non-parametric statistic (Mann–Whitney *U*) was used to test for a difference between trained and untrained tutors on each item. The values of *U* were significant beyond the .05 level for the first five Behaviors in Table 1. Nonsignificant *U's* were calculated for Behaviors 6 and 7.

When displaying Behavior 1, the tutor talks with his pupils about things other than the lesson. (In the training program tutors were given several "conversation starters" which they could use, e.g., "Do you have any pets at home?") The trained tutors exhibited this behavior a total of twenty-seven times compared to only once among the untrained tutors.

When displaying Behavior 2, the tutor verbally confirms correct responses by his pupil. The trained tutors confirmed 98 percent of the time compared to 49 percent for the untrained tutors. The observers noted, however, that while the trained tutors varied the prescribed confirmation phrases ("That's right," "Very good," "Yes"). the untrained tutors replied mostly with "uh-huh" supplemented by an occasional "O.K."

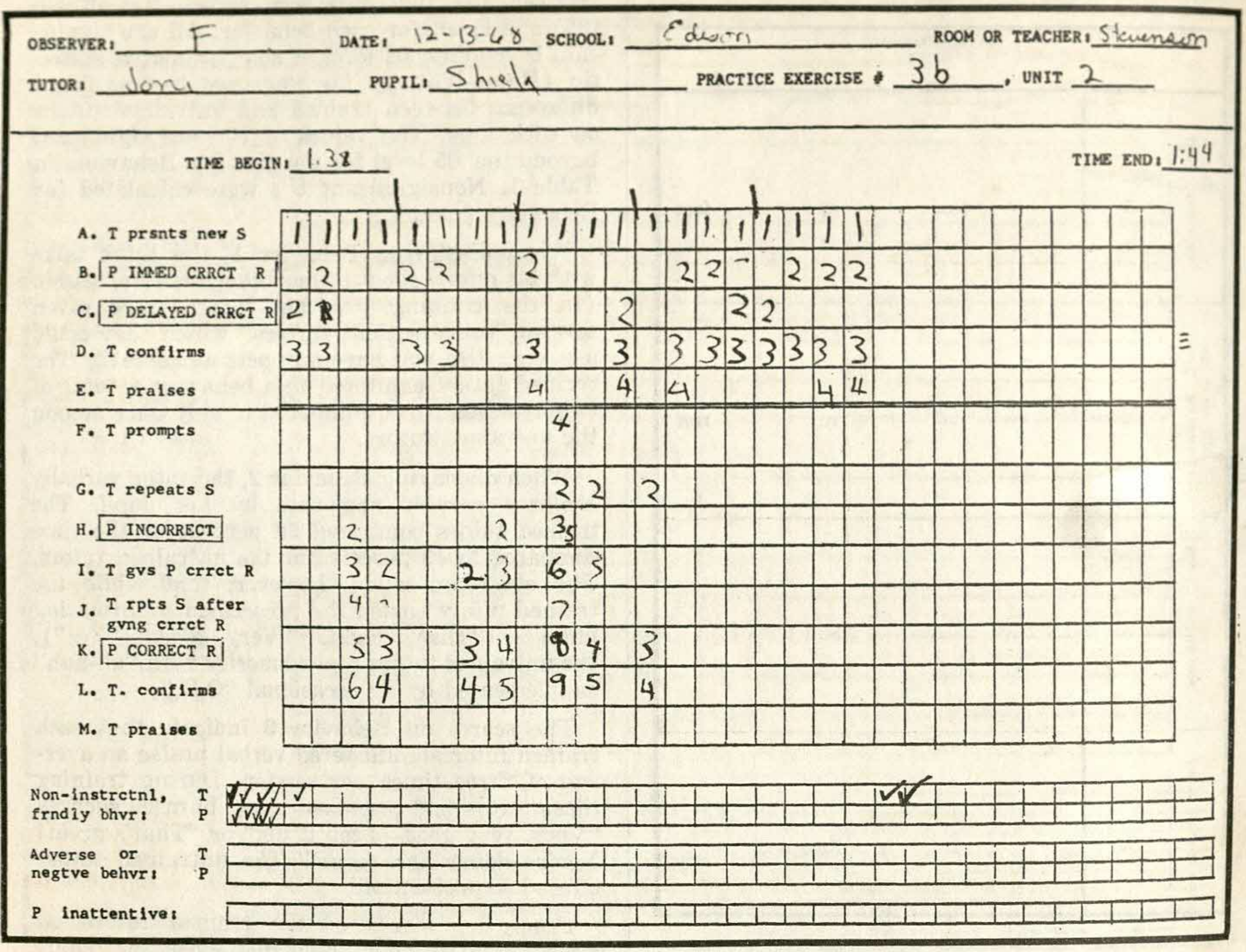
The scores on Behavior 3 indicate that each trained tutor administered verbal praise an average of three times per session. During training these tutors had practiced using phrases such as "Very, very good—keep it up," or "That's great! You're doing fine today." The untrained tutors offered no praise at all.

From the scores of the trained tutors on Behavior 4, it can be seen that when the pupils gave wrong responses, the tutors then told (or showed) them the correct response 85 percent of the time. The untrained tutors did this only 6 percent of the time (1 out of 17). Rather than tell the pupil the correct answer, untrained tutors would either repeat the question or try to elicit the correct response by prompting (see Behavior 7).

Behavior 5 states that, following instances of Behavior 4, the tutor should have his pupil make the response correctly before going on. For example, following an incorrect pointing response, the trained tutor might say, "You pointed to *man*. This word is *mad*. Now you do it. Point to *mad*." The trained tutors did this sixteen out of twenty-two times (75 percent).

If the pupil fails to respond, Behavior 6 states that the tutor should repeat the question or direction using different words. This objective was formulated on the notion that, if a child did not respond, he probably did not understand what was wanted. Trained tutors exhibited Behavior 6 only three out of twelve possible times (25 percent). Instead, they would treat nonresponses

Figure 3. Completed tutor observation scale for one ofthe six trained tutors



OBSERVER: F DATE: 12-13-68 SCHOOL: Edison ROOM OR TEACHER: Stevenson

TUTOR: Joni PUPIL: Sheila PRACTICE EXERCISE # 3b , UNIT 2

TIME BEGIN: 1:38

TIME END: 1:44

A. T prsnts new S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1										
B. P IMMED CRRCT R		2			2	2			2					2	2	−		2	2	2										
C. P DELAYED CRRCT R	2	R										2				2	2													
D. T confirms	3	3			3	3			3			3		3	3	3	3	3	3	3										
E. T praises									4			4		4					4	4										
F. T prompts										4																				
G. T repeats S										2	2		2																	
H. P INCORRECT			2					2		3 5																				
I. T gvs P crrct R			3	2			2	3		6	3																			
J. T rpts S after gvng crrct R			4							7																				
K. P CORRECT R			5	3			3	4		8	4		3																	
L. T. confirms			6	4			4	5		9	5		4																	
M. T praises																														

Non-instrctnl, frndly bhvr:	T	✓✓	✓✓	✓																			✓✓												
	P	✓✓✓	✓																																
Adverse or negtve behvr:	T																																		
	P																																		
P inattentive:																																			

as incorrect responses and simply tell the pupil the correct answer. Untrained tutors repeated the question following nonresponses four out of eighteen times (22 percent). Usually in this situation the untrained tutor would sit silently in hopes that the pupil would finally respond.

During developmental work with four tutors during the summer, it was noted that they had a tendency to prompt when the pupil was wrong or did not respond. It was felt, however, that the quality of their prompts was not high, i.e., their prompts usually failed to elicit the correct response. Thus, in the training program that evolved, tutors were told not to give hints, but instead to display Behaviors 4, 5, and 6. The results obtained with the Tutor Observation Scale seem to confirm the appropriateness of the no-prompt rule. The scores for Behavior 7 show that trained tutors prompted a total of only three times. Three of the untrained tutors, however, prompted a total of twenty-five times. Although not reported in Table 1, it is important to note that only seven of the twenty-five prompts by untrained tutors led directly to a correct response—a success rate of only 28 percent.

Conclusion

It was found that the instructional behaviors of trained and untrained tutors are quite different. The instructional behaviors implied by the objectives of the training program are based on established psychological principles. If it is desired that tutors behave according to these principles, then they should be trained. Unfortunately, very few tutorial studies currently appearing in the literature mention the training of tutors.

Table 1.—Instructional Behaviors of Trained and Untrained Fifth-grade Tutors Working with Kindergarten Children

	Trained Tutors								Untrained Tutors								Correlation of Observer Agreement (Spearman's *Rho*)
Tutor Behaviors	1	2	3	4	5	6	T	$\overline{X}$ or %	1	2	3	4	5	6	T	$\overline{X}$ or %	
1. Tutor engages pupil in non-instructional, friendly conversation.	1	1	1	7	7	10	27	4.5	1	0	0	0	0	0	1	0.2	.81
2. Tutor verbally confirms correct pupil responses.	12/13	17/17	8/8	13/13	17/17	11/12	78/80	98%	10/14	9/18	7/7	9/11	3/15	3/19	41/84	49%	.93
3. Tutor praises the pupil.	0	2	0	4	4	8	18	3.0	0	0	0	0	0	0	0	0.0	1.00
4. Tutor tells or shows pupil the correct response when pupil is incorrect.	2/2	1/2	6/7	3/4	3/3	7/8	22/26	85%	0/2	0/2	0/3	0/3	1/7	0/0	1/17	6%	.90
5. Tutor, after displaying Behavior 4, then elicits correct pupil response before going on.	2/2	0/1	4/6	2/3	2/3	6/7	16/22	75%	0/0	0/0	0/0	0/0	1/1	0/0	1/1	100%	.98
6. Tutor, following non-response to his initial question or direction, repeats in different words.	0/2	0/1	1/5	2/4	0/0	0/0	3/12	25%	1/4	0/0	3/9	0/5	0/0	0/0	4/18	22%	.98
7. Tutor avoids attempting to elicit correct response by prompting.	1	2	0	0	0	0	3	0.5	8	0	6	11	0	0	25	4.2	.85

FOOTNOTE

1. Paper presented at the annual spring conference of the California Educational Research Association, Los Angeles, March 14–15, 1969. For further data, see The Development of a Tutorial Program for Kindergarten Reading Instruction, TR. 26, May 1970, Southwest Regional Laboratory for Educational Research and Development, 11300 La Cienega Boulevard, Inglewood, California 90304.

Forthcoming Articles

Noffsinger, Thomas; Lake County Regional Program for Handicapped Children, Mentor, Ohio and Virginia Dobbs; Pupil Personnel Director Metropolitan Nashville Public Schools, Tennessee
"TEACHING ARITHMETIC TO EDUCABLE MENTALLY RETARDED CHILDREN (REVIEW)"

Watermann, Alan S.; Rensselaer Polytechnic Institute, Troy, New York and Caroline K. Waterman; State University of New York at Albany
"THE RELATIONSHIP BETWEEN EGO IDENTITY STATUS AND SATISFACTION WITH COLLEGE"

Weigel, Richard G. and Fred McKinney; Colorado State University, Fort Collins
"SELF-ORIENTED SENTENCE COMPLETION RESPONSES AND REPORTED PERSONAL BIOGRAPHICAL DATA"

Wiesner, Carol; University of Connecticut, Glastonbury
"A COMPARISON OF THE EFFECTIVENESS OF DISCOVERY VERSUS DIDACTIC METHODS AND TEACHER-GUIDED VERSUS INDEPENDENT PROCEDURES IN PRINCIPLE LEARNING"

Zastrow, Charles H.; University of Wisconsin, Madison
"CHEATING AMONG COLLEGE GRADUATE STUDENTS"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Interaction Effects of Sex of Mother's Child and Child's Reading Performance on the Mother's Evaluation of the School

OREN GLICK, University of Puget Sound

ABSTRACT

A recent project conducted within the Kansas City, Missouri, school district involved mothers in a multi-faceted effort to enhance the social-emotional and academic growthof third-year pupils who were having difficulty in learning to read.[1] It is commonly assumed that "parental support" of the school or positive parent–school relationships are conditions facilitating academic progress in the child. Accordingly, one of the intended outcomes of the project efforts was the "improvement" of parent evaluations of the school. The present paper reports the analysis of pretest data pertaining to the assumption of a relationship between parent evaluations of the school and academic success in the child. Do mothers whose children are having difficulty in learning to read evaluate the school less favorably than mothers of children not having difficulty?

This Kansas City, Missouri project was conducted in two primary schools located in middle-lower-middle white communities. All mothers who had third-year children scoring below the 3.0 grade equivalent score on all three portions of the Metropolitan Achievement Test of Reading were included in the sample. A random sample of 25 mothers whose third-year children performed at or above the reading criterion were also included. The original sample consisted of 112 mothers, twenty of whom were eliminated because they had moved after selection or could not be reached or failed to provide complete data. The final sample consisted of twelve mothers of boys and twelve mothers of girls who performed at or above the reading criterion and forty-one mothers of boys and twenty-seven mothers of girls who performed below the criterion.

Procedures. The mothers were interviewed individually in their homes by a team of six interviewers[2] the last 2 weeks of October, 1967. The interview included an administration of Cantril's Self Anchoring Scale (1) adapted to the school setting. This procedure yielded qualitative information by asking the respondent to specify first those characteristics of the "best of all possible schools" and secondly the characteristics of the "worst of all possible schools." The focus of the interview then shifted to an 11-point scale with "0" representing the "worst of all possible schools" and "10" the "best of all possible schools." The scale was presented on a full-size page with the following instructions:

> *Here is a picture of a ladder. Suppose we say that the top of the ladder (pointing) represents the best possible school for you and the bottom (pointing) represents the worst possible school for you.*
>
> *(A) Where on the ladder (moving finger up and down ladder) would you say your school stands at the present time?*
>
> *(B) Where on the ladder would you say it stood 2 or 3 years ago?*
>
> *(C) Where on the ladder do think it will be 2 or 3 years from now?*

The retrospective and projected evaluations yield, from the respondents' subjective reference, an evaluation extended in time and might be interpreted in terms of a general optimistic or pessimistic stance assumed by the respondent *vis-à-vis* the school.

Analysis. A School (2) by Sex of the Respondent's Child (2) by Above versus Below on the Reading Criterion (2) by Past-Present-Future (3) unweighted means analysis of variance (4: 374–379) was performed on the school ratings. The major point of interest in the analysis was to determine whether mothers whose children performed satisfactorily on the reading test would evaluate the school more favorably than mothers whose children performed unsatisfactorily. The additional factors in the analysis were included to suggest limits on the generalizability of the expected finding.

Results

The main effect for the Reading Criterion factor in the analysis of variance did not approach significance. The mean for mothers whose children performed above the criterion was 7.42; that for below criterion mothers was 7.48. The difference is negligible and in the reverse of the expected direction. However, the interaction of this factor with the Sex of the Mother's Child did approach significance ($p < .10$). For mothers of girls the means of those performing above and below the reading criterion were 8.15 and 7.50 respectively. The corresponding means for mothers of boys were 6.68 and 7.47. Thus there does appear to be a relationship between the mother's evaluations of the school and the academic performance of the child. It is positive, as expected, in the case of mothers of girls, but surprisingly, negative in the case of mothers of boys.

For mothers whose children fell below the reading criterion, sex of the child made no difference in the mother's rating. But for mothers of children performing above the criterion, mothers of girls evaluated the school much more favorably than mothers of boys. Of the four conditions, mothers of boys who performed above the reading criterion provided the *lowest* mean ratings. Inspection of the data for the two schools separately showed the same pattern of differences at both schools.

The main effect for Time was highly significant ($p < .001$). The mean ratings for "past," "present," and "future" were 6.89, 7.34, and 8.12 respectively. Thus a basic optimism, or maybe only wishful thinking, in the evaluations of the schools comes through strongly in these data. The present is better than the past, and the future appears even more promising.

Discussion

How does a mother accommodate to the unpleasant knowledge that her child is apparently performing relatively poorly in school? Three major alternatives are available to her in the attribution of responsibility for poor performance. (1) The responsibility may reside in the child and his characteristics. (2) The responsibility may reside in the behavior patterns of the mother and her relationship with the child. (3) Responsibility may reside in the characteristics of the school and its operations. It may be noted that the mother is in a poor position for obtaining sound and reliable information on the actual sources of difficulty, conditions which facilitate the resolution of the question in accordance with cognitive dissonance notions (2).

Consider first the likely resolution in the case of the mothers of boys. It is generally acknowledged that the world of school is more a girl's world than a boy's world. The dispositions of girls seem to fit better with the demands and requirements imposed by public school education than do the dispositions of boys. Thus it may be less surprising to a parent when a boy is having difficulty in school than when a girl is. Mothers of boys may already have gained some appreciation of the difficulties in socializing boys, and respond to the boy's unsatisfactory performance by seeking identification with the school as a supportive institution in the socialization of her child. Consequently, judgments of the quality of the school may become inflated and responsibility for the child's difficulties will likely be attributed to the child and his characteristics.

When a girl is having difficulty in school, support is more readily available for attributing responsibility to characteristics of the school and its operations. This option is compatible with the expectation that girls are "supposed" to do well in school. Thus, when a girl is having difficulty in school the mother is more likely to come to the support of the girl and attribute responsibility for failure to the school.

According to the above reasoning, mothers of boys who are having academic difficulty will differ from mothers of girls in their attribution of responsibility for the difficulty. This hypothesis was checked on fourteen mothers involved in the project by their responses to the following 11-point scale:

With respect to any difficulties that your child may be having in school, to what extent do you think the causes of those difficulties come from the child and his characteristics or from the school and the way it operates?

About half and half

Entirely in the child *5 4 3 2 1 0 1 2 3 4 5* *Entirely in the school*

The prediction is that the responses of mothers of boys will tend to fall to the left of the responses of mothers of girls.

The distribution of responses appears in Table 1. The results are predominantly in accordance with the hypothesis. Of the four mothers of girls, none responded to the left of the "half-and-half" point of the scale; of the ten mothers of boys, eight responded to the left of the "half-and-half" point. Scored left to right from 1 to 11, (6 being the "half-and-half" point) the mean value for mothers of girls was 7.00; for mothers of boys it was 4.25. Calculation of the randomization test for two independent samples (3) yielded $t = 2.34$ (one-tailed $p < .025$). Thus the hypothesis that mothers of boys and girls differ in their attribution of responsibility for the school difficulties of their children is clearly supported.

Table 1.—Distribution of Mothers of Boys and Girls on the Attribution of Responsibility Rating Scale

		Rating Scale										
		5	4	3	2	1	0	1	2	3	4	5
Mothers of	Boys	0	0	4	2	1	3	0	0	0	0	0
	Girls	0	0	0	0	0	3	0	0	0	1	0

These results should stimulate more definitive research on the processes involved in the attribution of responsibility for school failure by parents. The ideas here were stimulated by differences that were relatively weak statistically, though common to both schools in the study. They generated a hypothesis that was strongly supported, albeit with a small number of cases. The pattern of results might have been more pronounced if the effects of other children of the same mothers, of either sex, doing well or poorly in school, could have been controlled in the design.

FOOTNOTES

1. The Parent Volunteer Project was directed by Oren Glick and co-directed by John Hartley and Carl Thompson, principals of the two schools involved with the assistance of Mrs. Sandra Paap. The Institute for Community Studies, Kansas City, Missouri, generously provided financial assistance. A final report is in preparation.
2. Appreciation is expressed to Margaret Dart, Director of Volunteer Services at the Western Missouri Mental Health Center, for acquiring volunteer interviewers, most of whom were highly experienced and trained. Our thanks, also, to Marthann Luzader who scheduled and coordinated the interview operation.

REFERENCES

1. Cantril, H., *The Pattern of Human Concerns*, Quinn and Bodin Co., Inc., Rahway, New Jersey, 1965.
2. Festinger, L., *A Theory of Cognitive Dissonance*, Stanford University Press, Stanford, California, 1957.
3. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, McGraw–Hill Book Company, Inc., New York, 1956.
4. Winer, B. J., *Statistical Principles in Experimental Design*, McGraw–Hill, Inc., New York, 1962.

BOOK REVIEWS—Continued from Page 114

and examined, the report has relevance for anyone who has either general interest in beginning reading or specific interest in the application of i.t.a. It is broad in scope, carefully presented, and, in the reviewer's opinion, unbiased and restrained.

The results of a survey regarding the extent of i.t.a. use in England and Wales is given in the very brief Part One. Of the 163 local school authorities queried, 158 responded; of these 140 reported that i.t.a. was used with primary pupils in one or more schools. The main finding was that i.t.a. was used in 1,554 schools, 9.2 percent of the total from the 158 responding authorities. Only about 2 percent of the schools had returned to the use of traditional orthography (t.o.) for beginning reading after using i.t.a. They returned almost exclusively for administrative rather than instructional reasons.

In Part Two, the major portion of the report, the results of interviews regarding the impact and acceptance of i.t.a. are presented. Interviews were conducted in three main situations: (1) visits to local education authorities where the views of local inspectors and teachers were gathered; (2) appointments with people outside the local educational establishment and attendance at conferences and lectures; and (3) three group discussions held to fill in gaps in the overall coverage. Most of the individual interviews consisted of five stages: (1) an explanation of the project, (2) a request for general impressions of i.t.a., (3) questions designed to extend the discussion on various points, (4) factual questions regarding background, and (5) concluding questions designed to elicit summaries of general impressions. The general nature of the interview evidence is reflected in the advantages and disadvantages of i.t.a. as seen by primary teachers and by visitors to the schools and in the results from schools engaged in experiments with i.t.a.

The primary teachers, or infant teachers as they are called in Great Britain, noted three main groups of advantages for i.t.a. The main advantage related to the feeling that children learned to read and write earlier and more easily. Teachers felt both reading and writing were simpler with i.t.a., that children experienced more pleasure and satisfaction and less frustration, and that the number of non-readers was reduced. Second, children's ability and eagerness to tackle unknown words was considered a major advantage. Related to this facet were the observations that children could read a variety of books outside the basic reading program and that there were advantageous effects on performance in content areas. Finally, advantages related to attitudes, behavior, and social relationships were noted. Children displayed more confidence and independence in their school work with resulting improvement in general behavior, attitudes, and relationships. Teachers exhibited greater willingness to work together and greater personal satisfaction. Of the three main disadvantages noted, two were more often termed "dangers" than "disadvantages." These were the disadvantages encountered by teachers of transition classes; the danger of allowing children to transfer to traditional reading too soon, and the danger of permitting pupils to read beyond their comprehension. Some teachers, of course, pointed out the limitations in quality and variety of i.t.a. reading materials; but, again, this is less a disadvantage of i.t.a. method than a basic scarcity at a given point in time.

Visitors to the schools—including local education authority officials and advisers, Her Majesty's Inspectors, and University and College of Education lecturers—had generally favorable impressions of i.t.a., but they expressed certain "doubts and queries." Among the most

(Continued on page 138)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Computer-Assisted Planning and Scheduling of Individualized Programs of Study in Science and Mathematics at the Secondary Level

SISTER MARIJANE WERNER, Kent State University, Kent, Ohio

ABSTRACT

In order to explore the potential contribution that one form of critical path analysis, namely Program Evaluation and Review Technique (PERT), could make to a systematic ordering of interrelated units of study in science and mathematics, data were obtained from 312 experienced high school teachers of science and/or mathematics with respect to suggested time estimates together with a sequential ordering of topics to be studied by individual students majoring in these specialized areas of study. With data collected, a modified form of the IBM PERT Program # 10.3.006, run on the IBM 1620—40K computer, produced output within a matter of a few minutes. This information facilitated the drawing of the network model—a sketch which enables all parties concerned to see the critical path(s) which indicates points in scheduling time requiring special attention. Based upon an analysis of the statistical and descriptive data gathered in the study, the results showed that knowledge of PERT concepts and principles makes for more efficient preplanning of goals and objectives than might otherwise be the case. Used concurrently and consistently with other aspects of curriculum planning, PERT reveals interdependencies and problem areas which are neither obvious nor well defined by conventional planning and scheduling methods.

THE PIECEMEAL character of our educational system today is the essence of machine technology which has fragmented, centralized, and created a kind of superficiality in its pattern of human relationships. McLuhan (6) claims that the essence of automation technology is just the opposite. It is integral, interrelated, and comprehensive. Education is now under increasing pressure to acquire the depth and interrelation that are indispensable in the world of electronics. Wilhelm (9) believes that tomorrow's curriculum will be less fragmented than today's. The former is going to be attuned much better to the growth needs of the individual. A cursory examination of educational periodicals reveals that a major thrust in educational research today is that of individually prescribed instruction. Faced with the problem of educating young people for jobs which do not exist and for professions which cannot be described, Brown (2) emphasizes the need for a type of flexible education which is designed to produce adaptable individuals. U.S. Commissioner of Education Harold Howe II told the 1968 annual conference of the American Association of School Administrators that individualized instruction should be promoted in every school in the nation.

In view of the increasing number of nationwide efforts to individualize instruction at all levels, an efficient procedural technique needs to be found for handling the increasing complexity of the many new planning tasks which are implicit in attempting to provide individual programs of study for all students. The task of planning thousands of such programs by contemporary methods becomes an impossible task. Therefore, it is imperative that some computer-assisted technique for planning and scheduling be found in order to enable educators to perform these new tasks efficiently and quickly.

Problem

The purpose of this study was to explore the potential of a systems approach to the planning and scheduling of a program of study for individual students and to determine the extent to which a form of critical path analysis, namely PERT, could contribute to a systematic ordering of interrelated units of study in mathematics and science at the secondary level. The present investigation grew out of basic research with PERT conducted in the fields of business and industry. To the writer's knowledge, the potentialities of PERT network planning have not been investigated at the secondary education level. It is hoped that the reader will appreciate the limitations of an exploratory study. To expect the findings to become instantly practical would be premature at this time. For those unfamiliar with the principles of PERT, the basic ones are supplied.

Basic Principles of PERT. The PERT network has been described as a pictorial representation of all the work necessary to reach a terminal objective. It is a graphic portrayal of a series of accomplishments necessary to complete a program or project. Because the plan is visibly mapped out, potential or actual trouble spots can be pinpointed. Attention can be focused on these areas, and solutions can be worked out before the program/project gets underway. Used concurrently and consistently with other aspects of curriculum planning, PERT shows great promise in the field of educational planning.

A PERT network is concerned with specific accomplishments. Each accomplishment, in this study the completion of a subject matter unit, is called an event and is represented on the network by a circle. Each time duration—time needed to learn the subject matter unit—is called an activity. A line with an arrowhead represents an activity on the network. An activity line always connects two events. The event–activity relationship is the fundamental component of the PERT network.

The technique of applying PERT consists of first establishing a systematic breakdown of events required to complete the program. The following steps are suggested:

1. Define the end objective precisely.
2. Define all significant events that are precedent to the end objective. The "team" responsible for this task must make sure no important events have been omitted.
3. Define all significant events that are immediate predecessors of the events that preceded the end objective. Work back, a level at a time, making sure that all significant precedent events are established.
4. Make sure that all events except the beginning and end ones have at least one connection on each end.

PERT associates an expected time (t_e) with each activity. In order to determine in advance what this time is likely to be, it is necessary to estimate. The estimating procedure is the cornerstone of the PERT technique. Persons who have actually performed the activity many times before are asked for three time estimates: an optimistic time (a)—in this study, the number of hours it would take to learn a specific unit if everything went off exceptionally well. This is an unrealistic estimate to the extent that this estimate can be expected to occur in approximately one out of a hundred cases; a most likely time (m)—the time which, in terms of the estimator's past experience, this activity is most likely to take in view of the circumstances which are expected to exist; and a pessimistic time (b)—an estimate of the longest time an activity would require under the most adverse conditions. Again, this is an unrealistic estimate to the extent that it is expected to occur in approximately one case out of a hundred.

While the actual time the activity can take place may fall anywhere between the two extreme limits set by the optimistic time and the pessimistic time, the chances of it occurring at either of these two extremes are relatively small. On the other hand, the chances of it occurring on, or near, the most likely time are high. Thus, if the probability of any particular time occurring is plotted graphically, the assumption is that the shape of the curve is a "beta" distribution, with a range equal to time between the optimistic and pessimistic estimates, its peak situated at the most likely time, and having a standard deviation of one-sixth of the range. In lieu of the distribution curve, the formula, $t_e = \frac{a + 4m + b}{6}$ is used.

The t_e's for the entire network are calculated by the computer. This enables the systems analyst to determine the critical path through the network. The critical path is defined as the longest path in time from the starting event to the end objective event. It follows from this that all other events and activities must lie on paths which are shorter. When the scheduled completion time (T_S) for the total network is equal to or greater than the time required by the critical path, then all other paths contain slack time—time to spare. Computer programs often refer to "slack" as "float." Positive slack is an important sector in the planning phase of the total project.

In order to measure the amount of slack existing in any point on the network, it is necessary to determine the earliest expected time (T_E) and the latest allowable time (T_L) for each event. The former term is defined to be the sum of t_e's for the activities along the longest path leading to the starting event up to the event in question. T_L is determined by adding the expected times for activities on the longest path backward from the end objective event to the event in question and by subtracting this sum from the T_S for the end objective event. Slack, then, is the difference between the T_L's and the T_E's. Slack represents flexibility and is a range of time over which an activity can take place without delaying the accomplishment of the end objective event. The IBM 1620 computer PERT program #10.3.006 is capable of calculating this information for hundreds of events in a matter of a few minutes. A listing of the computer output need not be transferred to the network.

These are the basic principles of PERT. The network and the mathematical analysis of the network have several possible advantages over contemporary methods of educational planning: (1) It enforces a certain discipline in the overall logic of decision-making activities. (2) It is an ideal communications device. The advantages of having all members of a curriculum committee discuss and argue out details of relationships, inclusion of significant events, estimations of the three time elements required by PERT, and a sequential ordering of events, with a network before them while they are planning the complete program, cannot be overemphasized. (3) When properly developed, a network is an admirable means of searching out all the intricacies and interrelationships among elements of a projected operation to whatever depth of detail is believed necessary for its proper administration.

Network development and critical path analysis reveal interdependencies and problem areas which are neither obvious nor well-defined by conventional planning methods. To the person not familiar with the approach, it may appear too structured, too mechanical, and too definitive. It should be emphasized that the educator should never be bound to a particular system. This is especially true if the system inflicts undesirable rigidity on the operation. However, clarification of objectives and thorough planning should never be construed to be synonymous with rigidity and inflexibility.

Procedures

In order to obtain worthy, realistic data which would serve as a basis for the construction of a PERT network in an educational setting, a group of 600 high school teachers of science and/or mathematics was invited to participate in this study. Based on recommendations made by PERT experts, Alsaker (1), Cook (3), and Evarts (4), those responsible for coordinating actual activities should appraise the three time estimates required by PERT, and these people should assign predecessor events for each event on the network. Preliminary planning of the network should be undertaken by those who will eventually carry out the plan portrayed by the network. These suggestions dictated the necessity for a purposive selection, rather than random selection, of a sample. It was deemed advisable to require a minimum of 3 years teaching experience in science and/or mathematics at the secondary level. More than 75 percent of the sample population had attended NSF Institutes. Qualified teachers were selected from all geographical sectors of the United States.

Based upon recommendations from groups of high school teachers of mathematics and/or science with whom the writer consulted on many occasions, 160 units were chosen as the essential content of a 4-year program of interrelated topics in science and mathematics. It will be noted that the final selection of content reveals a "middle-of-the-way" approach. Some of the innovative curriculum changes were considered too recent to enable the teachers to contribute the kind of information required for this study.

Each participating teacher received a packet of materials consisting of a letter of introduction, a brief description of the project, basic PERT symbols and definitions of basic terms, a list of the 160 subject matter units arranged alphabetically and numbered consecutively, and ten reporting forms, each form bearing the name of one of the units from the list of units and requesting the three time estimates required by PERT together with information about predecessor units. The teachers were asked to complete as many of the ten forms as they felt competent to complete. In this way the investigator hoped to obtain at least ten responses per unit from as many different teachers.

The instrument was field tested before the packet was produced in quantity. Minor revisions were made before the final form of the packet was prepared and mailed to the participating teachers. Several hundred in excess of the 1,600 responses sought as a minimum by the investigator were received as the 312 teachers supplied 2,052 responses. As each form was returned, the three time estimates and the predecessor units suggested were tallied on a 160 × 160 grid. A unit was considered to be a necessary predecessor of another successor unit if two-thirds of the teachers responding to a given successor unit indicated that it should be so. By adhering to this criterion, the probability of irresolvable loops in the PERT network was considerably reduced. In this present study, no loops were encountered.

In order that this information could be machine processed, a standard data processing card for each of the 2,052 unit responses was keypunched. A second set of cards containing specific information about predecessor units was keypunched. These two sets of cards served as input data. A program written especially for these data directed the IBM 1620 (40K) computer to calculate the arithmetic mean for each of the three time estimates of each of the 160 units and to compile the information from both sets of cards onto a single deck of cards. The output from this operation was in the form of a set of machine-punched cards, 982 of them. This set then became the input data along with Part I of the IBM PERT Program—♯10. 3.006. The computer reprinted the mean of the three time estimates and calculated the earliest start times (ES), earliest finish times (EF), latest start times (LS), latest finish times (LF), total float (TF), and free float (FF). Once again the output data was in the form of machine-punched cards, arranged in the consecutive numerical order previously assigned to the alphabetical listing of the units. This output data had to be rearranged according to an ascending order of EF's, thus enabling the investigator to trace out the critical path on the network and, also, to determine a sequential ordering of those units which did not appear on the critical path. The IBM sorter performed this operation. An IBM accounting machine listed all computer output, thus enabling any interested party to have a printout of the results of the computer calculations.

In an attempt to determine how consistently the teachers were in their estimates of the three times required by PERT, a split-half reliability coefficient was found by correlating mean estimates of the first half of the total number of estimates for each unit with mean estimates of the remaining half of the total number of estimates for each unit. In the case of an odd number of responses, the middle response was averaged with the first half of the estimates. A program for this operation was available on the IBM 360 computer; therefore, the raw score formula of the Pearson product-moment coefficient of correlation was used.

According to Tate (8), the reliabilty coefficient computed from split-halves time estimates tends to underestimate reliability. Application of the Spearman–Brown correction formula adjusts the estimate to a more satisfactory reliability coefficient.

$$r_{XY} = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2]}} \quad \text{Tate (8:135)}$$

A split-half reliability coefficient for each of the three time estimates was determined.

$$r_1 = \frac{2r_{1/2}}{1 + r_{1/2}} \quad \text{Tate (8:187)}$$

Results

The product of innovative efforts in the field of education is frequently misunderstood. Silberman (7) who pioneered the field of systems approach to educational problems observed that failure, trial, and error are inherent in a truly innovative situation. New systems are made to work; they are seldom satisfactory on the first try. Design efforts differ from efforts to demonstrate and evaluate. Controls are restrictive by their very nature. Restrictions reduce the freedom to explore.

With these thoughts in mind, the results of this study are discussed. The findings represent a composite picture of the thinking of 312 high school teachers of mathematics and/or science. The respondents were interested in and enthusiastic about the project, but they expressed some apprehension about the merit and worth of their time estimates. The split-half reliability coefficient for each of the three time estimates after applying the Spearman–Brown correction Formula are summarized in Table 1.

Table 1.—Reliability Coefficients for the Three Time Estimates Required by PERT

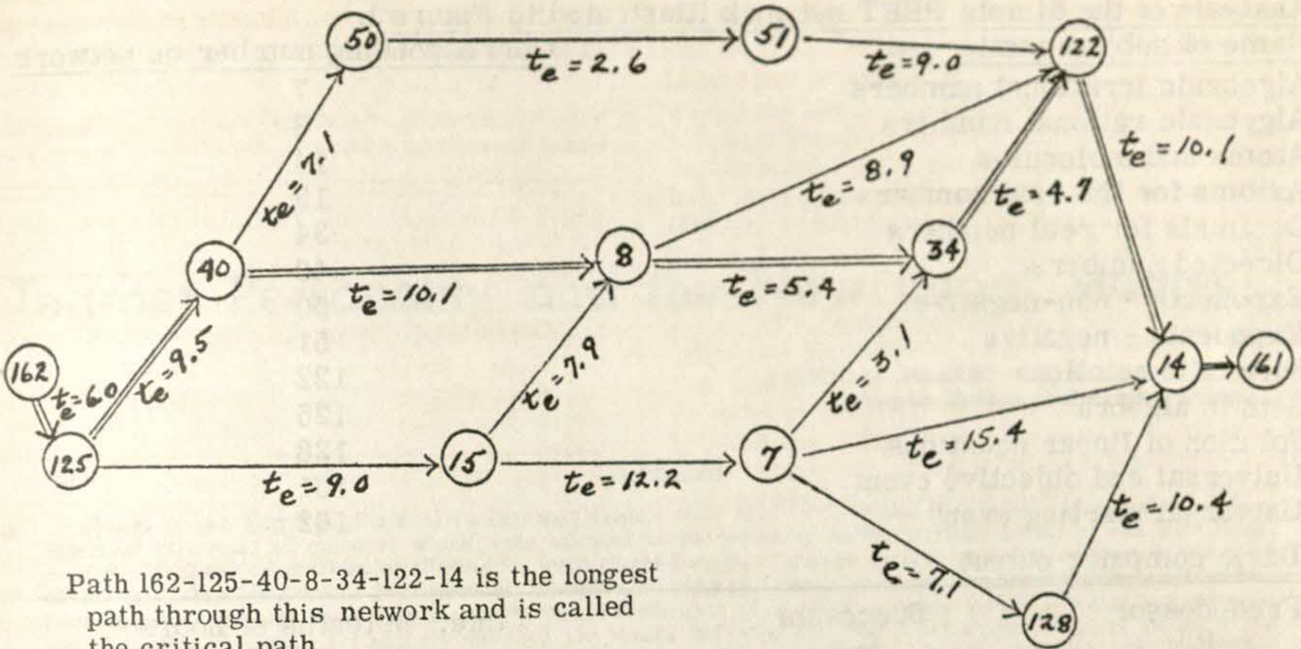
Time Estimate	Reliability Coefficient
Optimistic	0.799
Most Likely	0.793
Pessimistic	0.796

The data in Table 1 point out the fact that teacher respondents were very consistent in estimating the three time estimates required by PERT. These coefficients are considerably higher than the investigator had previously anticipated. The three-way estimate appears to result in a greater degree of honesty and accuracy in time forecasting.

It was also noted that the sum of the t_e's of all 160 units was 1,531 hours. This estimate of total time for the entire program would average about 191 hours of instruction, including classroom and laboratory experiences, for each of the eight conventional course offerings which comprised the total project. When one considers the large number of teacher respondents and the geographical expanse represented by this group, the total average t_e compares very favorably with actual time spent in learning much the same content during eight regular school semesters.

In accord with recommendations made by PERT practitioners that data be supplied by persons responsible for planning and implementing the project, the network for this study was drawn exclusively from data supplied by 312 high school teachers of mathematics and/or science. Space will permit only a simplified portion (see Figure 1) of the total network. The encircled numeral refers to the name of the unit in the alphabetical listing. The t_e's refer to the expected number of study hours necessary to complete a given unit before proceeding to the next unit. In order to lay the ground work for further investigation of this systems approach, it was necessary to work backwards. To improve the effectiveness of the technique, one must use the task analysis approach recommended by Gagné (5). Decide on the terminal objective, then step by step map out the prerequisite learnings and the proper sequences for events in the program of individually prescribed instruction.

FIGURE 1. A Simple PERT Network Illustrating Its Potential as an Aid to Educational Planning of Curriculum at the Secondary Level



Path 162-125-40-8-34-122-14 is the longest path through this network and is called the critical path.

For an analysis of the PERT network illustrated above, see Appendix.

Discussion

The PERT network has several advantages which augur well for educational planning: (1) It produces a clear outline of the activities that must be undertaken and of the events that must occur before an end objective is reached. The network provides a clear and readily understandable description of a plan before it is undertaken. (2) It is an ideal communications medium. The advantage of having teams of educators discuss and argue out details of relationships, inclusion of significant events, estimated times and the like, with a network plan before them while they are doing it, cannot be overemphasized. Time and energy can be recovered before they are wasted. (3) Many alternate successor events provide a flexibility which is desirable in providing for educational needs of individual students. (4) The network provides for periodic check points for purposes of evaluation and modification of an ongoing program of study. Enrichment material or additional personnel may alter the time factor. The network serves as a kind of progress chart which will need continual updating.

Judging from the dearth of interrelationships among subject matter units in science and mathematics revealed in this study, the plea for classroom teachers to break away from compartmentalized curriculum appears to have gone unheeded. This deficiency was graphically portrayed by the network when it was color coded.

Based upon the findings from this research, PERT planning and scheduling has greater potential for self-paced rather than group-paced programs of instruction. PERT planning is like swimming, repeated practice perfects the art, with team effort preferred to individual endeavor.

REFERENCES

1. Alsaker, E. T., *The Basic Technique: Network Analysis*, in Blood, J. W. (ed.), *PERT: A New Management Planning and Control Technique*, American Marketing Association Report, New York, 1962, pp. 37–80.
2. Brown, B. F., *The Nongraded High School*, Prentice-Hall, Englewood Cliffs, New Jersey, 1963.
3. Cook, D. L., *Program Evaluation and Review Technique: Applications in Education*, U. S. Government Printing Office, Washington, D. C., 1966 (OE - 12024 Cooperative Research Monograph No. 17).
4. Evarts, H. F., *Introduction to PERT*, Allyn and Bacon, Boston, 1964.
5. Gagné, R. M., *The Conditions of Learning*, Holt, Rinehart, and Winston, New York, 1965.
6. McLuhan, M., "The Medium is the Message," *NEA Journal*, 56: (no. 7) 24–27, 1967.
7. Silberman, H. F., "The Effect of Educational Research on Classroom Instruction," *Harvard Educational Review*, 36: (no. 4) 295–296, 1966.
8. Tate, M. W., *Statistics in Education and Psychology*, Macmillan, New York, 1965.
9. Wilhelm, F. T., "Curriculum Improvement in a New Key," *National Association of Secondary School Principals Bulletin*, 50: (no. 8) 79–86, 1966.

APPENDIX

Analysis of the Simple PERT network illustrated in Figure 1.

Name of subject matter unit	Corresponding number on network
Algebraic irrational numbers	7
Algebraic rational numbers	8
Atoms and molecules	14
Axioms for the real numbers	15
Decimals for real numbers	34
Directed numbers	40
Exponents - non-negative	50
Exponents - negative	51
Scientific notation	122
Sets in algebra	125
Solution of linear equations	128
Universal end objective event	161
Universal starting event	162

Basic computer output

Predecessor unit	Successor unit	Time: in tenths of hours			
		$\underline{a}$	$\underline{m}$	$\underline{b}$	$\underline{t}_e$
162	125*	4.1	5.5	9.8	6.0
125	15	5.6	8.9	13.0	9.0
125	40*	6.4	9.5	13.0	9.5
40	50	4.0	7.0	10.5	7.1
40	8*	6.4	10.1	14.1	10.1
15	8	5.7	7.9	10.2	7.9
15	7	8.9	12.0	16.0	12.2
50	51	1.5	2.6	3.8	2.6
8	34*	3.3	5.4	7.7	5.4
51	122	5.8	8.9	12.8	9.0
8	122	5.4	8.7	13.0	8.9
34	122*	2.4	4.3	8.6	4.7
7	34	1.5	3.0	5.2	3.1
7	128	5.3	6.9	9.7	7.1
7	14	11.5	15.9	17.3	15.4
128	14	7.9	10.3	13.4	10.4
122	14*	6.9	10.2	13.3	10.1
14	161	Not calculated because this is only a portion of the entire network			

* indicates the critical path

Unit Number	T_E	T_L	Unit Number	T_E	T_L
7	27.2	27.9	40	15.5	15.5
8	25.6	25.6	50	22.6	24.1
14	54.8	45.8	51	25.2	26.1
15	15.0	15.7	122	35.7	35.7
34	31.0	31.0	125	6.0	6.0
			128	34.3	35.4

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Teacher Personality and the New Social Studies

JAMES R. ROBERTSON and JOHN D. HAAS
Carnegie–Mellon University, Pittsburgh

ABSTRACT

Scales of the Edwards' Personal Preference Schedule were used to determine if selected social studies teachers possessed attributes of openness which were assumed to be necessary to the implementation of the New Social Studies. Teachers' perceptions of themselves were checked against student perceptions of these same teachers by use of the Robertson Student Perception of Teacher Attitude Scale (an instrument based on Ann Roe's categorizations of home climates). Data were treated by use of coefficients of correlation, comparison of F-ratios, and comparison of t-ratios. Although not entirely unmixed, the results indicated that the social studies teachers studied could handle the heat and pressure generated in open climates which are necessary in implementing the New Social Studies.

THE NEW social studies curriculum projects began in the late 1950's. With the advent of Project Social Studies the movement received added financial aid from the United States Office of Education. Currently, these curriculum projects are reaching their final stages. Like their forerunners in science and mathematics, the new social studies projects are *not* merely developing new and more efficient ways to look at subject matter; nor are they merely suggesting new courses and new materials. The social studies projects seem to assume that teachers will have open attitudes toward classroom discussion; that they will not be inclined to impose their ideas on the students; that classroom interaction is to be enhanced; and that teachers will be accepting of increased student responses.

The teacher qualities necessary to implementation of the "new social studies" have been variously described by leading proponents of this curriculum movement. To Fenton, teachers must choose styles and strategies from a continuum which proceeds from exposition to directed discussion to inquiry. Choices are based on objectives of instruction. There is little doubt that objectives and materials from the "new social studies" projects reviewed by Fenton prefer the inquiry or discovery end of the continuum (5). Teachers employing an inquiry mode or style will need to be "loose" and able to react openly to student responses—to roll with the ebb and flow of classroom dialogue.

To Massialas and Cox (10), the teacher is mediator of an open classroom which serves as a forum for inquiry. Students must be *allowed* to discover or rediscover concepts and generalizations which are the cement of disciplined knowledge.

The interplay of value positions and an environment for reflective thinking are concerns of Oliver and Shaver (12) and Hunt and Metcalf (7). Here, the teacher must learn to be comfortable and accepting in a climate of controversy—over public issues, value commitments, closed problematical issues, and the process of reflection.

Scales of the Edwards' Personal Preference Schedule (EPPS) get at these and other dimensions of teacher personality that could provide insights into a teacher's ability to use the new project products and approaches. The assumption is that the new social studies is not likely to be successfully implemented by teachers who have high needs for order or for deference to authority; or who cannot tolerate the "heat and pressure" of open climates. The purpose of this study was to determine whether social studies teachers were "ready for and amenable to" the implementation of the new social studies.

Specifically, the study sought answers to the following questions: (1) In terms of the items and scales of the EPPS, what manifest need patterns do social studies teachers exemplify? That is, do social studies teachers perceive themselves as secure, independent, worthy, self-enhancing, open, and accepting individuals? (2) How do these manifest need patterns differ from the norms of the general population on the Edwards' Personal Preference Schedule? (3) Do students' perceptions of their teachers corroborate the teachers' perceptions of themselves? And, finally,

(4) What implications can be drawn from the answers to the preceding three questions for the implementation of the new social studies?

Samples

The teacher sample for the study was drawn in Utah from the Ogden City Schools, the Davis School District, the Weber School District, and from the Social Science Experienced Teacher Fellowship Program at Utah State University. There were sixty-two Utah social studies teachers and fifteen social studies teachers in the Utah State University fellowship program. Because the Edwards' Personal Preference Schedule (EPPS) provides separate norms for women and men, and because of the low incidence of women social studies teachers, the sample was limited to male social studies teachers—a total of seventy-seven teachers.

Six high schools and six junior high schools from the northern Utah area were represented in the study. Of the seventy-seven social studies teachers in these twelve secondary schools, sixty-two participated in the study. Of the twenty-four social studies teachers in the Utah State University fellowship program, fifteen participated in the study.

The student sample was drawn from the classrooms of the sixty-two social studies teachers from the three school districts in northern Utah. This student group totaled approximately 3,700. No teacher's class was used which had been formed on the basis of any criterion of ability grouping (e. g. Advance Placement History). In general, one could say that the social studies classes used represented the typical courses found in a social studies curriculum in the nation—American history, civics, geography, economics, world history, psychology, sociology, and American problems.

Instruments

The instruments used in gathering data on the social studies teachers and their students were the Edwards' Personal Preference Schedule and the Robertson Student Perception of Teacher Attitude Scale. The EPPS is designed to measure fifteen manifest needs: Achievement, Deference, Order, Exhibition, Autonomy, Affiliation, Intraception, Succorance, Dominance, Abasement, Nurturance, Change, Endurance, Heterosexuality, and Aggression. Each of these single-word manifest needs are further elaborated and defined by Edwards (4).

The popularity of the EPPS seems to be due to its interesting variables and non-threatening items. It has been widely used to compare the need patterns of various professional occupational groups such as nurses, pharmacists, engineers, and naval officers. The EPPS also seemed to be a promising instrument for providing a prediction of the teacher personality need patterns requisite to use of inquiry approaches in the new social studies. Reported results in the literature suggest that one should expect high scores for social studies teachers in Deference, Order, and Endurance, and one might expect low preferences for Autonomy, Succorance, Exhibition, and Nurturance (2).

To confirm the teachers' perceptions of their attitudes, a second instrument was developed and administered to the students of the Utah teachers cooperating in the study. The Robertson instrument (Robertson Student Perception of Teacher Attitude Scale) was based on a similar instrument devised by Anderson and Dipboye (1). The Anderson–Dipboye instrument was in turn based on Ann Roe's categorizations of home environments: Rejecting, Demanding, Casual, and Accepting. Roe (13) provides descriptions of each of these four types of environments:

> *Rejecting:* A minimum or no affection is shown. There is persistent neglect. The environment is cold and hostile. The child is ridiculed for his inadequacies and problems.
>
> *Demanding:* The home is cold, materialistic, status seeking, authoritarian. Love is purchased by behavior. There is high punitiveness. The parents do not try to find out what the child is thinking; they tell him what to think.
>
> *Casual:* The parents pay some attention to the child and are mildly affectionate when they do. The parents do not think of him or plan for him much; they take him as part of the general situation. The parents are easy-going with a few rules; however, they do not make much effort to enforce those they have.
>
> *Accepting:* The parents give the child warm affection and try to help him with projects that are important to him. The child knows that he is accepted and wanted. Rules exist, but they are for the safety and growth of the child.

The Robertson instrument consists of six situations common to public school classrooms: poor performance or achievement, undesirable behavior, smoking, difference of point of view, scholastic difficulty, and improper dress. Each situation is presented in narrative form and followed by a forced-choice set of four possible responses based on the four Roe home climates. Scores were obtained by assigning three points to an Accepting response, two to Casual response, one to Demanding response, and zero to Rejecting response. While the instrument has construct validity based on the works of Roe and Anderson and Dipboye, its relationship to the EPPS was expected to provide further evidence of validity as well as to corroborate teachers' perceptions of themselves.

Treatments of Data

The EPPS was administered to the seventy-seven teachers participating in the study, while the Robertson instrument was administered to the 3,700 students. Fifteen t-ratios were computed between the means of the Utah teachers on the scales of the EPPS and the adult norms provided by the manual for the EPPS; and between the Utah teachers' means and the means of the teachers in the Utah State University fellowship program.

Coefficients of correlation were then calculated between the fifteen scales of the EPPS and the scores obtained from the students on the Robertson instrument. A further check was obtained by comparing F-ratios (between the standard deviations of the Utah teachers and the standard deviations provided by the EPPS norms and between the standard deviations of the Utah teachers and the standard deviations of the USU fellowship program teachers). Finally, the reliability of the Robertson scale was estimated with the Kuder–Richardson Formula 21.

In the analysis of the data, a confidence level of .05 was prescribed as sufficient to reject the following hypotheses: (1) There will be no difference between the norms established for the adult population on the means of the fifteen manifest need scales of the EPPS and the means of the selected Utah social studies teachers for the same scales. (2) There will be no difference between the EPPS means for the selected social studies teachers and the means of the social studies teachers in the Utah State University Fellowship program. (3) Coefficients of correlation between the EPPS manifest need scores of the Utah social studies teachers and the measured perceptions of their students by use of the Robertson scale will not be significant at the prescribed level of confidence.

Findings

The t-ratios for comparing the means of the selected Utah teachers and adult norms provided by the EPPS manual are presented in Table 1. The means for the Utah teachers were higher than those of the national norms at the .01 or .05 level for the needs of Exhibition, Intraception, Dominance, and Heterosexuality. Conversely, the Utah teachers means for the needs of Order, Abasement, Nurturance, and Endurance were all lower than the national norms at the .01 level. For the remaining needs (Achievement, Autonomy, Deference, Affiliation, Succorance, Change, and Aggression) there were no significant differences between the Utah teachers and the norm group.

Table 1.—Comparison of Means of Utah Social Studies Teachers with Adult Male Norms for EPPS

Need	Norm Group		Utah teachers		t
	Mean	Standard deviation	Mean	Standard deviation	
	N = 4031		N = 62		
Achievement	14.79	4.14	16.00	4.60	1.49
Deference	14.19	3.91	14.32	4.16	.17
Order	14.69	4.87	11.85	4.32	3.58**
Exhibition	12.75	3.99	14.37	4.03	3.16**
Autonomy	14.02	4.38	15.00	3.68	1.33
Affiliation	14.51	4.32	13.67	3.90	1.12
Intraception	14.18	4.42	16.34	3.78	2.90**
Succorance	10.78	4.71	9.64	4.80	1.37
Dominance	14.50	5.27	18.54	4.83	4.82**
Abasement	14.59	5.13	11.47	4.64	3.78**
Nurturance	15.67	4.97	12.60	4.10	3.95**
Change	13.87	4.76	14.56	4.74	.84
Endurance	16.97	4.90	14.10	4.61	3.51**
Heterosexuality	11.21	7.70	13.47	6.21	2.36*
Aggression	13.06	4.60	12.56	4.00	.65

* Difference with the EPPS norms at the .05 level of significance.
** Difference with the EPPS norms at the .01 level of significance.

Thus, the manifest needs which characterize the selected Utah social studies teachers were a relatively high preference for Exhibition, Intraception, Dominance, and Heterosexuality; and a relatively low preference for Order, Abasement, Nuturance and Endurance. This pattern suggests that the Utah teachers have an interest in achieving social recognition, being regarded as the leader of the classroom, being regarded as being attractive to the opposite sex, being concerned for the feelings of others, and a willingness to place oneself in another's position. Also, these teachers are not particularly interested in excessive detail and neatness in work, lack interest in involvement in another's difficulties, and have little interest in work requiring much patience and endurance.

The pattern found for the Utah social studies teachers is quite different from the image of the teacher described in the studies of Jackson and Guba (8), Johnson (9), Merrill (11) and Gray (6) in which teachers' needs were shown to be consistently high for Order, Deference, and Endurance, and consistently low for Succorance, Autonomy, and Exhibition. Only the needs of Affiliation and Nurturance (among those attitudes believed to be desirable for teachers) proved to be low in interest for the Utah social studies teachers.

The t-ratios obtained between the means of the selected Utah teachers on the fifteen EPPS scales and the means of the Utah State University fellowship program teachers for those same scales yielded the results depicted in Table 2. With only four significant differences, one could say that the Utah social studies teachers in the study are remarkably like the social studies teachers in the USU fellowship program.

Table 3 lists the coefficients of correlation between the teachers' scores for the fifteen scales of the EPPS and the scores derived from the students' ratings on the Robertson scale. For thirteen of the fifteen scales of the EPPS no significant relationships obtained. However, for Intraception, Aggression, and Affiliation the correlations were of sufficient magnitude to indicate a significant relationship existed between these scales and the Robertson instrument.

From the descriptions provided by Edwards, the three manifest needs of Intraception, Affiliation, and Aggression seem to approach most nearly those attitudes described by Roe for her climates. It might be expected, therefore, that a combination of the scores for these three needs

Table 2.—Comparison of EPPS Means of Utah Social Studies Teachers with Utah State University Social Science Institute Teachers

Need	USU institute teachers		Utah teachers		t
	Mean	Standard deviation	Mean	Standard deviation	
	N = 15		N = 62		
Achievement	12.93	2.39	16.00	4.60	2.77**
Deference	12.00	3.12	14.32	4.16	1.98*
Order	12.53	5.40	11.85	4.32	.48
Exhibition	13.13	4.45	14.37	4.03	.95
Autonomy	16.87	3.97	15.00	3.68	1.52
Affiliation	13.60	4.79	13.67	3.90	.05
Intraception	15.73	4.14	16.34	3.78	.48
Succorance	7.87	4.33	9.64	4.80	1.33
Dominance	17.67	4.65	18.54	4.83	.64
Abasement	12.53	4.32	11.47	4.64	.80
Nurturance	11.33	5.81	12.60	4.10	.88
Change	16.80	4.73	14.56	4.74	1.63
Endurance	14.27	4.13	14.10	4.61	.13
Heterosexuality	17.13	5.93	13.47	6.21	2.39*
Aggression	14.93	3.28	12.56	4.00	2.01*

* Difference with the institute teachers significant at the .05 level of significance.
** Difference with the institute teachers significant at the .01 level of significance.

Table 3.—Means, Standard Deviations, and Coefficients of Correlation between EPPS Variables and Robertson Pupil-Teacher Scale

Variables			
Achievement	16.00	4.60	.122
Deference	14.32	4.16	-.164
Order	11.85	4.32	-.128
Exhibition	14.37	4.03	.009
Autonomy	15.00	3.68	.097
Affiliation	13.67	3.90	.244
Intraception	16.34	3.78	.257*
Succorance	9.64	4.80	-.023
Dominance	18.54	4.83	.008
Abasement	11.47	4.64	-.169
Nurturance	12.60	4.10	.076
Change	14.56	4.74	.128
Endurance	14.10	4.61	-.162
Heterosexuality	13.47	6.21	-.027
Aggression	12.56	4.00	-.272*
Robertson Pupil-Teacher Scale	12.50	1.36	

*Significance at .05 level.
N = 62

would better relate to the accepting–rejecting continuum of the Robertson scale.

Combining the teachers' scores for Intraception and Affiliation and correlating this to the Robertson scale yielded an r of .306. When the teachers' scores for Intraception and Affiliation (the two scales most nearly approximating "Acceptance") were summed, and then the teachers' scores for Aggression (the need most nearly approximating "Rejection") were subtracted from that sum, this yielded a higher relationship. An r of .347 (p .10) obtained between this combination of EPPS variables and the Robertson scale.

Discussion

A key assumption underlying this study was that the new social studies curriculums will be best implemented by teachers who hold open, accepting, and understanding attitudes. This assumption is supported by writers in the field such as Cox and Cousins (3), Oliver and Shaver (12), Fenton (5), Hunt and Metcalf (7), Rogers (14), and Massialas and Cox (10). The descriptions presented by these writers of classroom climates which are believed to be necessary to the successful implementation of reflective inquiry are in marked contrast to the descriptions of classroom climates found in research literature. A review of the literature reveals that public school classrooms are characterized by "endemic" conflict between the achievement expectations of the school structure and the goal expectations of students. This conflict is aggravated by the tacit but powerful rating system for determining teacher merit on the basis of ability to maintain classroom control. Studies have revealed the consistent pattern of teacher needs for Deference, Order, and Endurance, and low preferences for the "more appropriate" needs of Intraception, Affiliation, Change, Nurturance, Autonomy, and Achievement. From such studies the judgment that teachers are like the stereotype advanced by the mass media (impotent, obsequious, patient to a fault, and socially inept) seems undeniable. The literature has indicated that even the most basic teacher personality requirements—needs for social service, for interest in children—seem to be lacking. That teachers like those described in the literature could implement the new social studies could seriously be questioned.

The results of the present study indicate that Utah social studies teachers do not see themselves like teachers described in the literature. That is, they do not see themselves as socially inept, conforming, deferring to authority, and long suffering. Indeed, comparisons with the Utah State University fellowship program teachers raise some doubt that social studies teachers generally match the stereotype. On the contrary, social studies teachers report strong social orientation without inhibiting feelings of guilt or inadequacy. Of course, the picture is not entirely unmixed. A high preference for exhibition indicates that a social studies teacher may be too interested in being the focal point of attention and thus detract from student initiated responses in the classroom. A teacher's need for dominance, for maintaining control of group situations, could also inhibit students from responding and reacting in class. Nevertheless, the data from the present study indicate that Utah social studies teachers possess need patterns which could enhance the implementation of the new social studies curriculums.

Can present social studies teachers deal with reflective inquiry in the classroom and handle the heat and pressure that characterize societal disputes which become the focus of classroom discussion? We conclude that the answer is a cautious "Yes!"

REFERENCES

1. Anderson, William F.; Dipboye, Walter, "Student Perception of Parent Attitude Scale, (Mimeo) Syracuse University, Syracuse, New York, 1962.
2. Buros, Oscar Krisen, *Sixth Mental Measurements Yearbook*, The Gryphon Press, Highland Park, New Jersey, 1965.
3. Cox, C. Benjamin; Cousins, Jack E., "Teaching Social Studies in Secondary Schools," in Massialas, Byron; Smith, Frederick R., (eds.), *New Challenges in the Social Studies*, Wadsworth Publishing Company, Belmont, California, 1965, pp. 101–119.
4. Edwards, Allen L., *Edwards Personal Preference Schedule* (Manual), The Psychological Corporation, New York, p. 27, 1959.
5. Fenton, Edwin, *The New Social Studies*, Holt, Rinehart, and Winston, New York, 1967.
6. Gray, J. T., "Needs and Values in Three Occupations," *Personnel and Guidance Journal*, 42:238–244, November, 1963.
7. Hunt, Maurice P.; Metcalf, Lawrence, *Teaching High School Social Studies*, Harper and Row, New York, 1968.
8. Jackson, P. W.; Guba, E. G., "The Need Structure of Inservice Teachers: An Occupational Analysis," *School Review*, 65:(2)176–192, 1957.
9. Johnson, Lynn Eric, "Personality Profiles of Teacher Graduates," unpublished master's thesis, University of Utah, Salt Lake City, 1959.
10. Massialas, Byron G.; Cox, C. Benjamin, *Inquiry in Social Studies*, McGraw–Hill Book Company, New York, 1966.
11. Merrill, Reed M., "Comparison of Education Students, Successful Science Teachers, and Educational Administrators on the Edwards Personal Preference Schedule," *The Journal of Educational Research*, 54: 38–40, October, 1960.
12. Oliver, Donald W.; Shaver, James P., *Teaching Public Issues in The High School*, Houghton–Mifflin Co., Boston, 1966.
13. Roe, Ann; Siegleman, Marvin, *The Origin of Interests*, American Personnel and Guidance Association, Washington, D. C., p. 7, 1964.
14. Rogers, Carl R., "Learning to Be Free," *National Education Association Journal*, 52:(3)28–30, 1963.

BOOK REVIEWS—Continued from Page 126

prominent doubts and queries were these: (1) children who use i.t.a. in school will be surrounded by t.o. outside of school; (2) children can, when properly taught, succeed in reading with t.o.; (3) the use of i.t.a. may lead to undue emphasis on phonics; (4) the use of i.t.a. may lead to stress on reading and neglect of other aspects of the curriculum; (5) children may have great difficulty "unlearning" when they transfer to t.o.; (6) children who transfer from i.t.a. to t.o. schools may suffer disadvantages; (7) quality and quantity of i.t.a. books is still limited; (8) i.t.a. does not help the slowest pupils. Prominent advantages noted by the visitors were these: (1) children learn to read earlier and more easily; (2) children soon become confident and self-reliant, and they read more books on their own; (3) there are fewer non-readers in the primary (infant) school; (4) written work is of increased quality and quantity; (5) the transition appears to present no major problem; (6) the use of i.t.a. leads to improvement in the standard of reading instruction, particularly among the poorer teachers.

The evidence from the schools involved in experiments with i.t.a. is best sumed up in the words of the authors of the report:

> Of the seven headmistresses of infant schools who were conducting their own experiments, none expressed disapproval of i.t.a., although a few minor disadvantages were pointed out. Their views may be summarized simply by noting that five approved of i.t.a. and intended to continue to use it, either for part or the whole of the school; two had not yet made up their minds; none said she disapproved of i.t.a. (p. 102).

In the final portion of the book, Part Three, seventeen independent investigations of i.t.a. in the United States and Great Britain are closely scrutinized, summarized, and evaluated. In the words of the authors of the report, "The general conclusion drawn from these various researches is that i.t.a. is a superior medium to t.o. in teaching young children to read, and that this advantage may be lost after transition" (p. 275). On the basis of all the evidence, which Wharburton and Southgate have gathered and examined with admirable competence and care, the conclusion seems justified.

Teachers who have used i.t.a. clearly see more advantages than disadvantages, and the existing research results favor i.t.a. as an *initial* teaching alphabet. But again, the authors of the report are themselves the first to offer a word of restraint:

> It would be unfortunate if the mainly favorable tone of this report was taken to imply that the use of i.t.a. for beginning reading with infants was the final and only solution. The experiments of the past six years have demonstrated that in many cases the usual ways in which children have been taught to read can be improved. We should continue to investigate how children learn to read and the most effective media, methods, materials and procedures for helping them (p. 284).

Meanwhile, what are the main implications of the report for the school staff that is contemplating the adoption of i.t.a.? Perhaps most basically, i.t.a. is likely to do no harm and teachers are likely to respond positively, so i.t.a. appears to be an innovation that can be adopted without fear of negative effects. The use of i.t.a. is likely to be accompanied by improved performance in reading and writing at the primary level. Whether the improvement is in fact due to the use of i.t.a. or to an accompanying Hawthorne effect is of little importance if the adoption of i.t.a. is predictably accompanied by improved early reading. But the question that remains is whether the added trouble and expense is justified by the short-term effect. The answer, at least for the present, can be given only in terms of the value placed upon early reading. In the opinion of the reviewer, the price is too high.

Finally, it should be clear that Wharburton and Southgate have taken a purely pragmatic approach in their investigation. Their concern was with the impact of i.t.a. in practice; they did not seek to question the foundations or the assumptions of i.t.a. Nor is this review the place to raise such questions, although the need to continue to raise them is urgently acknowledged.

Wayne Otto
The University of Wisconsin—Madison

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 3, November 1970)

Test Difficulty as a Factor in Achievement

JAMES R. DAINES
Stout State University, Menomonie, Wisconsin

ABSTRACT

The study investigated the effect of two testing approaches on informational achievement in a technical industrial education course. Two equated groups of students were administerd either a high difficulty or low difficulty form of unit achievement examinations which served as the experimental factor. Informational achievement was measured by pretest-posttest gain of each testing approach. Comparisons were made (1) between the total populations of the treatment groups, (2) of the ability levels between the treatments, and (3) of the ability levels within each treatment. The findings revealed that while significant gains were found in the informational achievement of both groups, the gains associated with the "low difficulty testing approach" were significantly higher than those of the "high difficulty testing approach."

WHILE testing as an evaluative device is a widely accepted and traditional aspect of education, the use of an examination as a learning instrument is often overlooked. Considering the time spent in the design, administration, and analysis of examinations, the instructional value of the testing process deserves attention in the teacher education program. Many professional education courses emphasize the mechanics of examination construction for evaluation purposes through various types of questions, grading formulas, and criteria for ranking. While these aspects serve a distinct objective and are necessary for reliable and valid evaluation, the function of utilizing the examination as an aid to learning should also be considered. The items included in the instrument should stimulate further learning by the student, so the question then becomes one of determining the type of examination item which best accomplishes the goal of encouraging self-education.

One aspect of examination preparation which concerns teachers is the relative difficulty of examination items. Opinions vary regarding the value of difficult versus easy examinations. Some teachers believe that motivation to learn new information is increased if the student is confronted by difficult examinations; other teachers prepare relatively easy evaluation devices, assuming that high difficulty will discourage student initiative. Neither assumption appears to be supported by sufficient experimental evidence, and the optimum examination difficulty and informational gain relationship has not been identified.

An investigation conducted at the University of Hawaii by Sax and Reade (3) found that examination difficulty does affect student achievement significantly; the experimenters reached the general conclusion that hard examinations motivate the bright students to greater achievement than do easier examinations, even though this effect is not maintained for the less-able class members. However, the conclusion was not supported by a recent study completed at Stout State University which investigated the relationship of test difficulty and student informational achievement in a technical industrial education course.

Method

High and low difficulty testing approaches were the treatment variables in the investigation conducted at Stout State University, Menomonie, Wisconsin, during the third quarter of the 1967-1968 school year. A 2-group experimental design was used in the investigation, with data gathered in two sections of power mechanics which were taught by the experimenter. Each section was composed of twenty-five students. The basic procedure followed during the experiment was to randomize the groups, pretest to verify equivalence, apply the treatment, and posttest to ascertain the treatment effect.

Initial status of the research population was ascertained through the analyses of five control factors: (1) scholastic aptitude, (2) informational achievement in the power mechanics area, (3) anxiety in a testing situation, (4) attitude toward the course, and (5) previous courses in the power mechanics area. A t-test for independent samples applied to the control factor data indicated that no significant differences existed between the total treatment groups. Analysis of the data by ability levels revealed only one significant difference, that of the scholastic aptitude of the high and low ability student groups identified within each treatment variable. High ability students were defined as any student in the top forty percent of the experimental groups as determined by the composite score of The American College Testing Program Examination. Low ability students were likewise identified by scores ranking in the lower forty percent of the students taking part in the study. Each ability group in the investigation included ten individuals.

Eight teacher-constructed unit achievement examinations of varied difficulty levels were employed as the experimental factor. One section was administered a series of examinations consisting of items with a difficulty of forty or less, and the second section in the study received examinations with items having a difficulty of sixty or more. Mean difficulty of the examinations is shown in Table 1. Similar administration, scoring, and discussion procedures were employed in each section. The 5-distractor, multiple choice items used in the achievement examinations were verified in regularly scheduled power mechanics sections during the three instructional terms that preceded the study; the verification sections were also taught by the experimenter. Analyses of item difficulty, item discriminating power, and response effectiveness were made. After each verification period, the individual items of the unit achievement examinations were analyzed and rewritten in an attempt to meet the criteria of discrimination and difficulty.

Table 1.—Mean Difficulty of Unit Achievement Examinations Administered throughout the Experimental Period

Examination	Examination Number			
Form	1	2	3	4
Low Difficulty	78.40	73.87	75.20	84.27
High Difficulty	27.73	28.00	26.93	26.67

During the verification terms, students enrolled in the course through the regular registration procedure; however, during the experimental period the registration procedure was modified to allow the students to be randomly selected for each group. The experimental groups were assigned either the high or low difficulty testing approach by coin toss. Throughout the experimental and verification periods, an attempt was made to provide similar instructional content, supplies and equipment, and physical facilities. Discussion of the examination items following the scoring and analysis procedure was structured similarly, and each student received a standard t-score indicating his achievement on each test. All grading was based on standard scores.

Table 2.—Difference between Mean Gains in Informational Achievement by Treatment Variable

Low Difficulty Approach	High Difficulty Approach
mean gain = 34.32	mean gain = 26.76

F-ratio = 1.54*
df = 48
t = 2.46**

*Not significant at the .05 level of confidence.
**Significant at the .05 level of confidence.

Data for the initial and final analysis of informational achievement was obtained from an instrument constructed and verified by the experimenter. The items used in the examination were verified in the three instructional terms preceding the study, and included true–false, multiple choice, and completion items. Each form of the examination was analyzed for item difficulty, item reliability, and overall test reliability. Point biserial correlation was used in the analysis of item reliability with Hoyt's analysis of variance formula (2) utilized to ascertain test reliability. The final form of the examination contained 120 items with a full range of difficulty and with acceptable reliability characteristics. The reliability coefficient for the complete examination when used in the posttest of the experiment was .88.

Results

Analyses of the final status of the total treatment population were conducted using t-tests for independent samples, while the ability subgroups were examined with the Mann-Whitney U-test. Equality of variances for the two treatment groups was tested using the F-ratio. Statistical significance was tested at the .05 level of confidence for all analyses. Results of the analysis of pretest and posttest scores indicated a significant gain in informational achievement in both treatment groups. A mean gain difference of 7.56

Table 3.—Informational Achievement Gain by Treatment Variable within Ability Groups

	Low Ability Group		High Ability Group	
Test Difficulty	Low	High	Low	High
Mean Gain	28.10	22.60	41.20	28.50
R	119.50	90.50	142.50	67.50
E(R)	105		105	
U*	35.5		12.5	

*A value of 23 or less indicates a significant difference at the .05 level of confidence.

points on the informational achievement examination was recorded between the two treatment group populations with the "low difficulty testing approach" producing the highest gain. Table 2 shows the mean gains for each approach.

Table 4.—Informational Achievement Gain by Ability Group within Treatment Variables

	Low Difficulty Approach		High Difficulty Approach	
Ability Group	Low	High	Low	High
Mean Gain	28.10	41.20	22.60	28.50
R	66.00	143.00	91.00	119.00
E(R)	105		105	
U*	12		36	

*A value of 23 or less indicates a significant difference at the .05 level of confidence.

Further analysis of the informational achievement data revealed the following information: (1) the high ability students in the "low difficulty testing approach" experienced a significantly higher gain in informational achievement than did the high ability students who received the high difficulty examinations; (2) the high ability students who experienced the "low difficulty testing approach" exhibited a statistically significant higher informational gain than did the less able students in the same group; and (3) there was no significant difference between the gains of the high and low ability groups on informational achievement with the "high difficulty testing approach." The following tables show the analyses of the mean gains in achievement by treatment variable within ability groups, and by ability groups within the treatment variables.

REFERENCES

1. Daines, James Richard, "The Effect of Test Difficulty on Informational Achievement in a Technical Industrial Education Course at the College Level," unpublished doctor's dissertation, University of Missouri, Columbia, 1968.
2. Hoyt, Cyril, "Test Reliability Estimated by Analysis of Variance," *Psychometrika*, 6:153–160, June, 1941.
3. Sax, Gilbert; Reade, Marybell, "Achievement as a Function of Test Difficulty Level," *American Educational Research Journal*, 1:22–25, January, 1964.

Research Briefs

NOTE ON THE UTILITY OF ADVANCED ORGANIZERS

IN HIS two texts on educational psychology, Ausubel (1, 2) presented limited evidence for the effectiveness of advanced organizers, which are, in his words,

> ". . . more complex and deliberately prepared sets of ideas which are presented to the learner in advance of the body of (meaningful) material to be learned, in order to insure that relevantly anchoring ideas will be available" (2: 145).

There are various kinds of advanced organizers, all of which are to be distinguished from the usual overview material that precedes text material, and all are expected to facilitate both learning and retention of the text material that follow them.

In this brief investigation, seventy-four university sophomores in an educational psychology course were assigned a reproduced copy of chapter 3 on meaningful learning from Ausubel and Robinson (2:50–74) and were directed to study overnight the material in preparation for an examination at the next class meeting. Approximately one half the students ($N = 38$) were given copies of the chapter that were preceded by material (2:43–45) which Ausubel and Robinson (2:146) called an advanced organizer.

On a 20-item multiple choice examination on the chapter no significant differences ($t = 0.26$, $p > .05$) in the number of questions correctly answered were found between the group with the organizer (mean = 12.48) and the group without the organizer (mean = 12.65). Students were asked to write a short essay on the issues and points they would need to consider for their pupils to meaningfully learn a mathematical proposition. They were scored for the number of correct and relevant ideas they cited, and again there were no significant differences ($t = 1.16$, $p > .05$) in the number of these between the group with the organizer (mean = 3.05) and without the organizer (mean = 2.60). However there were significant differences between the groups in the amount of time each reported it spent studying ($t = 2.04$, $p < .05$). Subjects with the organizer reported a mean of 107.16 minutes of study and those without the organizer reported a mean of 129.39 minutes of study. Whether subjects actually devoted that amount of time or whether they felt that much time was spent is not known, and is somewhat immaterial to the conclusion that subjects with the organizer had a little easier time (real or imagined) acquiring the same level of competence as those subjects without organizer and the added chore of reading of it. Overall the results indicate that arguments for construction of advanced organizers for classroom use are not impressive, and indicate as Stephens (3) has noted that few organizational variables in education are sufficiently robust to survive, undiminished in vigor, the transplant from the laboratory to the classroom.—*Frank B. Murray, University of Delaware; Joseph R. Jenkins, Newark Special School District.*

REFERENCES

1. Ausubel, D., *Educational Psychology: A Cognitive View*, Holt, Rinehart and Winston, New York, 1968.
2. Ausubel, D.; Robinson, F., *School Learning: An Introduction to Educational Psychology*, Holt, Rinehart and Winston, New York, 1969, p. 145.
3. Stephen, J., *The Process of Schooling*, Holt, Rinehart and Winston, New York, 1967.

FIELD NEWS

Willard E. North

Central Missouri State College

A REVISED EDITION OF *Tests in Print* is being prepared by the Institute of Mental Measurements under the editorship of Oscar K. Buros for publication in early 1972. The revision will attempt to list all educational and psychological tests published as separates in English-speaking countries.

In order to insure that recently published materials are included in the revised *Tests in Print*, test authors and publishers are urged to submit specimen sets of their new and revised tests if they have not already done so. Tests should be sent to the Institute of Mental Measurements, 220 Montgomery Street, Highland Park, New Jersey 08904.

* * * * * *

A MONOGRAPH TITLED *Technical Report of a Project to Develop Educational Cost-Effectiveness Models for New York State* has been published by the New York State Department of Education. It should be of value to those interested in cost-effectivenes studies as well as those interested in ways to use multiple regression analysis.

This study was intended to explore the feasibility of using data extracted from the Education Department's data files to develop multiple regression equations required for simulating the effects, on pupil achievement, of alternative expenditure configurations. The study was one of a series focused on the development of regression analysis procedures for converting empirical data into information for the educational planner and decision maker. Starting with 187 variables for about five hundred New York School Districts, this study: (a) identified internally correlated clusters of predictive variables; (b) identified internally correlated clusters of criterion variables; (c) revealed the degree of association between each cluster of predictive variables and a representative variable from each of the criterion clusters; (d) yielded correlations between most of the individual predictive variables and each of the representative criterion variables, and (e) produced a set of eight linear and eight curvilinear equations for estimating various dimensions of School District output.

A copy may be obtained, free of charge, by writing to David J. Irvine, the University of the State of New York, The State Education Department, Albany, New York 12224.

* * * * * *

THE EXCEPTIONAL PARENT, A NEW magazine, is now being published by the Psy-Ed Corporation. It will aim to provide practical help for the parents of children with disabilities, and will combine the knowledge of experts with day to day experiences of laymen. The magazine will deal with many issues that affect the exceptional child and will cover such topics as the role of the family, the nature and role of various professional groups with whom the family is apt to come in contact, and the ways in which certain aids can be helpful. Information will be easily understandable, practical as well as theoretical. The first magazine will also provide a means for parents to exchange ideas, share concerns, and discover new approaches to problems.

The founders and editors of the publication are Louis Klebanoff, Stanley Klein, and Maxwell Schleifer. The editors will be assisted by an Editorial Advisory Committee, which includes Burton Blatt, Director of Special Education and Rehabilitation, Syracuse University; Michael Dukakis, attorney and member of the Massachusetts House of Representatives; Gunnar Dybwad, Professor of Human Development, Brandeis University; Hubert Grossman, Director, Illinios State Pediatric Institute; Goodwin Katzen, Executive Director of Rockland County (New York) Center for the Physically Handicapped; Albert Murphy, Chairman of the Department of Special Education, Boston University; and Barbara-Jeanne Seabury, Director, National Research Center, Washington, D. C.

Charter subscriptions are $6.00 per year. Further information may be obtained by writing *The Exceptional Parent*, Box 45, Newtonville, Massachusetts 02160.

* * * * * *

BILINGUAL SCHOOLING IN THE UNITED STATES, a monograph which represents the first attempt at providing a comprehensive picture of bilingual schooling in this nation and other parts of the world, is now available. The 2-volume monograph, written by Theodore Andersson, Program Director, and Mildred Boyer, Coordinator of The Bilingual Education Learning System of the Southwest Educational Development Laboratory, traces bilingual schooling through its historical, developmental, and cultural perspectives to its current implication for society. It was written under contract with the U.S. Office of Education.

The authors made three extensive field trips to visit sites where bilingual programs are in progress. The first led through the Northwest as far north as Barrow, Alaska; the second was a tour of the Southwest and Hawaii; and the third was from Dade County, Florida, north through Washington, New York, New England, and into Canada.

Besides a detailed history of bilingual schooling, the monograph includes alternative concepts of bilingual schooling, sample curriculum models, an outline of needs as related to action and research, an extensive bibliography, and general discussion of the issues surrounding the subject. Although there are currently more than one hundred bilingual programs in operation in the public elementary schools, the authors believe that the current programs are inadequate in extent and sometimes in quality. According to the authors, the major difficulties seem to lie in the lack of public understanding about the goals of such programs; a lack of consensus among bilingual educators on appropriate approaches to such programs; and a great shortage of trained bilingual teachers.

Copies of the monograph may be obtained through the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. The price is $6.00 for the 2-volume set. Sets cannot be divided.

* * * * * *

THE ERIC CLEARING HOUSE on Teacher Education recently surveyed members of the education profession to determine the methods used to acquire information. A generalized profile of about three hundred correspondents indicates that, in seeking information, the user tends to go to his personal files, a library, his colleagues, and lastly to information service organizations such as ERIC/SP. One notable exception to this generalization is that those responding who were identified as educational research and development people at the college level, and those who were identified as educational personnel at the pre-college (K-12) level, tended to view the information service organization more favorably as an information source then did those identified as college instructors and administrators.

ERIC/SP correspondents have also given some opinions about the kind of document preferred. The original document is preferred as a source of information by most, abstracts and objective reviews or summaries of a number of related documents generally rank second; bibliographies and evaluative or interpretive summaries of a number of related documents were preferred by the smallest number of correspondents.

Regarding the use of ERIC tools, about half of the ERIC/SP correspondents have used Research in Education and ERIC/SP in newsletters and publications. Approximately half have access to Research in Education; about one-third, to Current Index to Journals in Education; less than half have access to an ERIC microfiche collection. A larger proportion of pre-college staff and Educational Research and Development people in college use microfiche more than did either college instructors or administrators. The number of correspondents who have found the ERIC Thesaurus useful in searching is about the same as the number who have never used it.

The information on user preference reported suggests that some attention should be given to making ERIC services more readily available to potential users and to diversify the output of clearing houses to correspond with the information needs of a number of different audiences.

* * * * * *

A JOINT INTERNATIONAL PROGRAM of Research, Training and Exchange of Educational Information is currently being worked out by five universities. The School of Education of the University of Wisconsin (UW) is working closely with schools at the American University of Beirut, Lebanon; the University of Ife, Ile-Ife Nigeria; the Universidad de Oriente Caracas, Venezuela; and Ahmadu Bello University, Zaria, Nigeria. They plan to develop and submit a joint proposal to the Ford Foundation for exchange programs extending over 3 to 5 years.

Professors Andreas M. Kazamias and Robert L. Koehl of the UW's Department of Educational Policy, the originators and coordinators of the project, envision a "circular or intermingling" exchange among the universities involving both personnel—graduate student and faculty—and research data information. The idea was originated 3 years ago when Professor Kazamias met with representatives from the Schools of Education, Agriculture, Engineering, and area studies and drew up a proposal to establish the "Institute for Comparative and Interdisciplinary studies in Educational Policy." One of the key hopes of the Institute was that a number of universities could freely exchange research information among themselves and that channels of communication could be open. At a recent international conference, representatives from the participating universities made commitments to the program and agreed to develop a joint funding proposal. Currently, individual universities are preparing their proposals for inclusion in the draft grant.

* * * * * *

THE TEACHERS CORPS WILL commit its entire funding in the next 5 years exclusively to the support of universities attempting a radically different approach to teacher education and to school systems which seek the use of the Corps to introduce their practices. According to Richard A. Graham, Director of the Teacher Corps, approximately $30 million will be committed during the first year of the program. This fundamentally different approach to training teachers is most commonly called "competency-based education." In essence, it means that teachers would no longer be certified because they have passed required courses but because they have acquired and demonstrated competencies in a wide spectrum of teaching methods and skills.

Teacher Corps grants will be used to support only those colleges, universities, and school districts, which seek to develop programs which are based on the principles of the Models of Elementary Teacher Education developed during the last 2 years by leading universities under grants from the National Center for Educational Research and Development (NCERD) and the Office of Education. Universities seeking new programs will not be asked to select any one of the NCERD models, but can use the principles common to these models as guides in developing competency-based programs of comparable scope.

The NCERD models were developed at Florida State University, Michigan State University, Syracuse University, University of Georgia, University of Massachusetts, University of Toledo, University of Wisconsin, University of Pittsburgh, Columbia University, and the Northwestern Regional Educational Laboratory of U.S. Office of Education. Seven universities are already developing Teaching Corps programs using these models as a guide. They are Oakland University at Pontiac, Michigan; Buffalo State University College, New York; San Francisco State College; Texas Southern University, Houston; Western Kentucky State College at Bowling Green; Jackson State College, Jackson, Mississippi; and Western Washington State College at Bellingham.

• • • • • •

GALTON INSTITUTE, PUBLISHERS of Perceptual–Cognitive Development (PCD), has launched a new publishing venture with a monograph titled *Modes of Biological Adaptation and A Role in Intellectual Functioning*, by W. Ragan Callaway, Jr., Associate Professor, Graduate School of Education, UCLA. The Callaway monograph is a synthesis of information previously reported in the literature. It brings together work done by biologists, neurologists, and geneticists on the molecular nature of gene action and relates it to the behavioral concepts such as the critical period hypothesis, Hebb's cell-assembly/phase-sequence theory, and Piaget's adaptation/assimilation model.

The newly inaugurated PCD monograph series is intended to serve as a countervailing force to the current trend toward reporting of fragmented and specialized research by providing a platform for the generalists, the relationists, and the creative synthesizer of other people's research who gives meaning to "old" findings within a discipline or across traditional disciplinary alliances.

Appropriate manuscripts are being accepted. Typewritten papers should follow the American Psychological Association style guide and should be submitted to: The Galton Institute, PCD Monograph Series, 319 South Robertson Boulevard, Beverly Hills, California, 90211. Single copies of the PCD monograph may be ordered from the Galton Institute at $4.95 per copy or for the 4-issue volume year, published on a non-scheduled basis, for $15.00

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

EUNICE ASKOV has her PhD from The University of Wisconsin and is in the Department of Curriculum and Instruction. WAYNE OTTO has his PhD from the University of Wisconsin and is currently Professor of Curriculum and Instruction and Principal Investigator, Wisconsin Research and Development Center for Cognitive Learning at The University of Wisconsin. He was formerly on the faculty of the University of Oregon and the University of Georgia. He is an Executive Editor of this Journal. WARREN ASKOV is an advanced graduate student at The University of Wisconsin.

JAMES R. DAINES is Professor of Power Technology at Stout State University, Menomonie, Wisconsin. He has taught in the industrial arts department at Mankato State College, Mankato, Minnesota, and also in the Michigan public schools. On several occasions he has assisted Science Research Associates in the construction of test items designed to evaluate vocational proficiency. He holds BS and MS degrees from Stout and the EdD from the University of Missouri, Columbia.

E. LARRY EBERLEIN, Assistant Professor, Department of Educational Psychology, University of Alberta, Edmonton, Canada, is a certified psychologist and is currently President of the Psychologist's Association of Alberta. He received his PhD from Michigan State University, BD from Garret Theological Seminary, JD from The University of Wisconsin Law School, and BS from The University of Wisconsin. His current research interests focus on personality theory, counseling, and group and social dynamics in education.

OREN W. GLICK received his PhD in social psychology from the University of Kansas. For 5 years he was a Research Associate at the Greater Kansas City Mental Health Foundation and the Institute for Community Studies in Kansas City, Missouri. At the same time he was involved primarily in research in the public schools. He is currently teaching child development and social psychology of education at Seattle University. Research interests include the effects of student-peer group conditions and processes on the achievement of educational objectives, evaluation of educational programs within a social systems framework, and techniques for inferring causality in non-experimental data.

LESTER S. GOLUB has his PhD from Stanford University. He was a Principal Investigator in the writing phase of the language arts project, at Wisconsin Center for Cognitive Learning and was a lecturer in English and curriculum and instruction at The University of Wisconsin. He was formerly on the faculty of the University of Nebraska and of San Jose State College. He recently joined the faculty of the University of Illinois, Chicago-Circle campus.

JOHN D. HAAS is presently Director of Teacher Education and Associate Professor of Social Science Education at Carnegie–Mellon University. He has taught in the public schools of Michigan and at Utah State University. He holds a BA from Hope College, Holland, Michigan, and MA and PhD degrees from the University of Michigan. JAMES R. ROBERTSON is a counselor at San Joaquin Delta College, Stockton, California. He has been a social studies teacher, counselor, and director of adult education in the Salt Lake City Schools and Davis County schools in Utah. He received his EdD from Utah State University and MS and BS from the University of Utah.

FRED C. NIEDERMEYER is a member of the professional staff at the Southwest Regional Laboratory for Educational Research and Development (SWRL) in Inglewood, California. SWRL is one of several U.S.O.E. regional laboratories. It develops validated classroom instructional programs for use in primary grades. Dr. Niedermeyer has been responsible for the development of tutorial programs. parent-assisted learning, and teacher training as they relate to a kindergarten reading curriculum.

SISTER MARIJANE WERNER, O. P. has her PhD from Boston College. She is currently an Assistant Professor in the College of Education at Kent State University, Kent, Ohio. Her teaching experiences range from grade 5 through graduate school. She has had National Science Foundation Grants for continued study in mathematics at Fordham University, the University of Akron, and Boston College. Dr. John A. Schmitt, Associate Dean of Graduate Studies, College of Education, Boston College, Chestnut Hill, Massachusetts, was chief consultant for the study reported here.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor Curriculum and Instruction, The University of Wisconsin, Madison.

PHILIP LAMBERT, Chairman, Department of Educational Psychology and Administration, The School of Education, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1970

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Assistant Professor of Special Education, Yeshiva University, New York.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English and Education, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and nstruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

ARL R. DOUGLASS, 4185 Pinion, Boulder, Colorado 80302.

ARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, alifornia 92802.

. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

LFRED S. LEWERENZ, Director Emeritus of Evaluation and esearch, 4107 Holly Knoll Drive, Los Angeles, California 0027.

ILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, lorida 33146.

. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 2601.

ESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

ARLETON W. WASHBURNE, Director Emeritus of Teacher ducation and the Division of Graduate Studies, Brooklyn ollege, 2248 Kent Street, Okemos, Michigan 4886

223
18. 3. 71
Library

File No
Bureau Ednl. Psy. Research

50th Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **DECEMBER 1970** **NUMBER 4**

ARTICLES

What Does Research Say About the Judging of Writing Ability?—*William McColly* 147

Cheating Among College Graduate Students—*Charles H. Zastrow* 157

A Longitudinal Study Examining Prediction of Doctoral Success: Grade Point Average as Criterion, or Graduation vs. Non-Graduation as Criterion—*John D. Williams, Steven D. Harlow, Del Gab* 161

The Relationship Between Ego Identity Status and Satisfaction with College—*Alan S. Waterman, Caroline K. Waterman* 165

Relations Between Grade Point Averages and Collegiate Course Grade Distributions—*William A. Mehrens, Bruce G. Rogers* 169

Effects of Certain Audio and Visual Presentation Sequences on Pupil Information Acquisition—*Armand J. Galfo* 172

Teaching Arithmetic to Educable Mentally Retarded Children (Review)—*Thomas Noffsinger, Virginia Dobbs* 177

Characteristics of Students Who Could Not Take and Schools Which Did Not Offer Driver Training—*Joe E. Shively, William Asher* 185

DEPARTMENTS

Among Those Present 192

Book Reviews—*Robert E. Clasen* 160

Field News—*Willard E. North* 190

Forthcoming Articles 168

It Says in the JER 146

Research Briefs—*James D. Evans* 171

New Books from the Publishers 156

(The Journal of Educational Research is indexed in Abstr.S.W., CIJE, CSPA, Current Contents, Ed. Adm. Abst., Educ. Ind., M.H.R.Ind., PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin.

Copyright, 1970 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, Supervisor of Editorial Services.

It Says in The JER...

WHILE REVIEWING THE research on the judging of writing ability, McColly found that opinion is divided over the comparative value of judging writing-test essays for general merit or for performance on each of several specific factors. The validity of a single rating is not questioned; the issue is whether or not an array of ratings may not be more discriminating (p. 147).

DO GRADUATE STUDENTS CHEAT? Zastrow found that about 40 percent of those in his study cheated on quizzes given as part of the experiment. They felt that pressure to obtain good grades was the main reason they had cheated in the past. There were no significant personality differences between cheaters and non-cheaters (p. 157).

IF A LONGITUDINAL STUDY is limited to a single university department, promising predictability can be found with the variable of graduated-not-graduated as the criterion for predicting doctoral success, say Williams, Harlow, and Gab. They found, generally, that the Graduate Record Examinations have the greatest predictive value, although other background variables do increase prediction somewhat (p. 163).

THE ASSOCIATION of crisis with college is an important factor in student dissatisfaction, suggest Waterman and Waterman. Students who were in the process of going through an identity crisis over occupational choice had the least favorable evaluations of their education while students who had never experienced a crisis but were committed to a vocational goal, held the most favorable attitudes. Students who went through the crisis over choice of major field while in college showed more negative attitudes than did students not experiencing such a crisis during college years (p. 165).

A MODEL WHICH MAY prove useful in predicting the effect of proposed classroom grading policy on GPA distribution prior to the actual implementation of that policy has been proposed by Mehrens and Rogers. The model was empirically tested and described in detail (p. 169).

SIMULTANEOUS SIGHT-SOUND presentation of material is superior to sight-sound separation. Galfo selected eighth and ninth grade students for an audio and visual presentation sequence experiment (p. 172).

RESEARCH RELATED TO ARITHMETIC ABILITY of mentally retarded children has evolved into a diversity of special learning areas. Noffsinger and Dobbs categorize eight special interest areas and highlight the unique contributions of each. They conclude that there is a need for continued research (p. 177).

FIFTY-NINE PERCENT of the schools studied by Shively and Asher offered driver education and 303 of the 798 students studied took the course, 151 could not take it because it was not offered. Investigation of the random sample of 861 high schools indicated that differences exist between those schools offering driver training and those not offering it (p. 185).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Editorial/Abstract

What Does Educational Research Say About the Judging of Writing Ability?

◆

The judging of the writing ability of students is a problem because of the great disagreement among English teachers as to the nature of good writing. Objective tests are of little help, since they are not measures of writing. Essay tests are the only valid measure, but to make them as reliable and as valid as possible requires that attention be given to a number of sources of error.

Typically the problem of error in writing tests is conceived of as more a matter of reliability than validity, and of arising from three general sources: students, readers, and topics. A number of other sources are important as well, however, and the aspect of validity is at least as important as reliability if not more so.

The importance of the fluctuation in performance of a single writer from task to task is perhaps exaggerated. If all other sources of error in a writing test are removed or controlled, the most logical strategy may be to accept a single performance as a valid one even if it is the writer's worst, since a test by definition is a decisive trial and *ipso facto* imposes the worst possible conditions. There are no relevant research findings here.

Reader competence, or knowledgeability, is crucial. Obtained coefficients for inter-reader reliability, which can be taken as one reflection of competence, indicate that other things being equal, professors of English are the most reliable readers. It has been demonstrated that competence can be increased through training and practice, especially in accelerating judging-rate.

Opinion is divided over the comparative value of judging writing-test essays simply for general merit or for performance on each of several specific factors. The validity of a single rating is not questioned; the issue is whether or not an array of ratings is not more discriminating. Research findings indicate that the controversy may be a *tour de force,* because ratings for specific factors tend to collapse entirely into a single rating through halo effect.

Research findings offer very little toward understanding the effects of several other important factors: number of points in a rating scale and the size or extent of a scale; differences in essay topics, especially in regard to structured versus unstructured topics, and testing time.

A factor which has often been ignored but which has been experimentally verified is that of test-paper appearance. The main aspect in appearance is handwriting, which has a systematic effect on readers that is obscured by reliability coefficients since the effect is correlated in the data.

Finally, the question whether computer "grading" of writing tests offers some hope is answered rather positively. The judgments of the computer so far published are unimpressive, but the hope lies in the fact that computer programs do work directly on the only valid basis for the judging of writing ability, written compositions.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

What Does Educational Research Say About the Judging of Writing Ability?[1]

WILLIAM McCOLLY
The University of South Carolina

TO judge the quality of verbal expression as expression seems to be a natural human trait and is perhaps nearly as old as speech itself. It is an element of the action in very early literary works, as in the *Odyssey;* Aristotle's *Rhetoric* and his *Poetics* reflect a compelling endemic interest in criticism. Today even the general student of literature who studies no criticism formally is nevertheless exposed to some of it through literary parody, but he may not be aware of the fact, since much literary parody is covert and very often English teachers themselves seem to be unaware of it.

Basis of the Problem

The problem at hand is educational, not literary. In the field of literary studies we do not say that there is the problem of criticism, we say nothing more emphatic than that criticism is an especially hard subject of inquiry and investigation. In the field of educational studies, on the other hand, there does exist the problem of the criticism or evaluation of student writing. The problem has its academic side, but its existence as a problem arises more from its professional than from its academic aspect. It is concerned not so much with the instruction of students as it is with the uses of the evaluation of that instruction: when a judgment of level of writing ability is a factor in the progress of students from one teacher to another, from one grade or ability level to another, or from high school to college. There is dramatic disagreement among the hundreds of thousands of English teachers as to what good writing is and is not, and also, many English teachers cannot be objective in judging student writing or in reporting their judgments. If all English teachers were unbiased and of uniform opinion, the problem would be merely technical, and we would have to overcome only the difficulties of large-scale communication. A teacher would somehow communicate, for example, with the Department of English of the state university that one particular student was an excellent writer and should be exempted from the freshman course and that another particular student was a terrible writer and should enter the non-credit remedial course, and the department would make these placements forthwith. But of course, to seriously consider such procedures is almost unthinkable.

The problem is not new. The College Entrance Examination Board was organized in 1900 for the purpose of helping the member institutions better assess the academic ability of their applicants. In 1901 two members of the Department of English at Harvard published a book about their freshman composition program, in which there is the following sketch of the students (6:2):

> *At one extreme of this class of Freshmen are the illiterate and inarticulate, who cannot distinguish a sentence from a phrase, or spell the simplest words. At the other are fairly mature writers, who need only to discard certain crudities and to gain variety and flexibility. Between these two extremes come many sorts and conditions of students.*

Today on the whole the freshmen at Harvard may be better writers than they were in 1901, but in general things have not improved relatively. Even in the schools and the colleges which have very able students, there are still the illiterate and inarticulate as well as the polished and fluent, with "many sorts and conditions" between.

However, because of the comparatively greater number of students, the problem of identifying these different levels of writing ability is today much more desperate.

Objective Tests versus Writing Tests

With the aim of coping with sizable aggregates of students, objective tests of writing ability, so called, were devised. There is a common notion that such objective tests are better than essay tests because they are more "reliable." In his chapter on essay examinations in the new edition of the American Council on Education's *Educational Measurement*, Coffman speaks to this point:

> *To reject a test of relatively low reliability which was really a sample of the underlying quality in favor of a more indirect measure of higher reliability might be to settle for lower validity for the measure* (4:64).

This is a cautious statement of the position normally taken toward such objective tests by scholars in the field of literature and language, that because the tests simply are not measures of writing, for the purpose of the judging of writing ability they therefore should be ignored. In essence that is the viewpoint adopted here.

Measures of Writing Ability: Essay Tests

The essay-type examination can be used for any of three different general purposes. Depending upon the form and substance of the topic and task, an essay can test (1) writing ability, (2) thinking ability, or (3) ability to put subject-matter mastery into the mode of written discourse. Of course, if the ability to write in a form other than the essay, such as in verse or fiction, is to be judged, the criterion performance must be in that particular form. But the kind of writing which is felt to be the basic responsibility of the schools and colleges to teach is ordinary, workaday, expository prose. Thus, nearly always the kind of writing test in which students actually write is thought of as an essay test.

What Makes Writing Tests Imperfect?

A writing test has two general aspects in which shortcomings render the test to some degree less than ideal: validity and reliability. It is often said that reliability is the more important of the two, on the basis of the popular notion that without reliability there is no value to validity. But really the inverse relationship is what is true. If a test is totally non-valid, but is nonetheless reliable in all respects, it may be worse than useless because of the likelihood of its misuse. Be that as it may, the scholarly literature that deals with writing tests shows more apparent concern with reliability than with validity. French classifies the "*sources of error in the scores on a composition test*" as (1) "*student error,*" (2) "*test error*" (the task and topic), (3) "*scale error*" (severity of reader judgment in general), and (4) "*reader disagreement*" (8:3–4). Braddock classifies these factors as (1) the "*writer variable,*" (2) the "*assignment variable,*" (3) the "*rater variable*" (within one rater), and (4) the between-rater variability, referred to as the "*colleague variable*" (2:6–15). These schemes are very similar and are representative of the thinking on the subject. Actually they can be collapsed into three general categories: students or writers, readers, and topics. These and several other sources of error and variability will be treated in the following discussion, which will have more emphasis on validity than this particular aspect has normally been given.

Writers

There are no research findings here, but the situation is not as bad as it sounds. Suppose that all other sources of error were to be removed or perfectly controlled; it would still be true, in Braddock's words (2:6), that we "*cannot be sure that the student is fully using his ability, is writing as well as he can,*" because

> *Something may be causing the student to write below his capacity: a case of the sniffles, a gasoline lawnmower outside the examination room, or some distracting personal concern.*

Suppose, again, that in an otherwise "error-free" and perfectly controlled writing test a particular student is indeed distracted by his respiratory condition and by noise, the question is, to what degree is his performance not valid? In the very act of giving a writing test we *ipso facto* create for the normal student writer not the best but the worst possible writing conditions. We create unavoidably the conditions which make of the test a decisive trial—which is the time-honored concept of a test. The testing of ability is a material and not an ideal process, and should be undertaken not in a test tube but in normal reality-settings. Thus, if a student during critical writing performances is distracted by the sniffles or by a lawnmower, who is to say that the measure of his performance should not reflect such distractibility? After all, distractions are very much a part of life, especially student life.

This author is well aware that some examiners and teachers will agree with the view just expressed and some will disagree. It should be pointed out that if one's opinion is that the student's "best" writing performance should be measured, one has a really insuperable task in decontaminating the test for certain factors, including the personal. And if one's opinion is that the student's normal, or typical performance should be measured, one has a terrible sampling problem.

Readers

The most important of the reader factors is that of *competence*. The principle which obtains goes almost without saying: the more competent the judges of essays are the more they will agree and the more valid will be their judgments. The kind of competence meant here is scholarship, or knowledgeability. Some normalcy of temperament is required as well for the demonstration of competence in essay-rating, but the incompetence which arises from the intrusion of neuroticism is generally a problem of administration, not scholarship.

In regard to the factor of competence there are findings of interest to report. At The University of Wisconsin during the academic year 1962–63 this author performed an experiment, which will henceforth be referred to as the Wisconsin High School Experiment, in which almost 1,200 essays were judged by sixteen English teachers: four college males, four college females, four high school males, and four high school females. The mean correlation coefficients for the essay-ratings by judge type were, respectively: .50, .39, .40, and .41. Obviously the last three do not differ from each other by very much, but a significance test based on Fisher z-transformations indicates that the coefficient for the college male teachers is higher than each of the other three at the .05 level. It so happened that the educational level of the college males was substantially higher than that of the other judge types (12).

In a large-scale, essay-rating study conducted by Diederich and two associates, three hundred expository essays (150 essays on each of two topics) written by college freshmen were judged by fifty-three male readers who were drawn from six occupations: English professors, professors of social science, professors of natural sciences, lawyers, businessmen, and writers and editors. The following reliability estimates are reported (7:14):

Median Correlations of Readers Within Fields:

English	.41
Lawyers	.38
Natural Science	.34
Social Science	.28
Writers–Editors	.26
Business	.22

Median Correlations With English:

English	.41
Lawyers	.39
Natural Science	.35
Social Science	.33
Writers–Editors	.30
Business	.30

It is not surprising to see that the professors of English reflected more unanimity than did the other groups. These data should be interpreted in light of the fact that one of the purposes of the study was actually to invite disagreement, *"to reveal the differences of opinion that prevail in uncontrolled grading"* (7:i).

It is plain that readers must be given the proper *training* and *orientation*, regardless of how knowledgeable they are. Two different approaches exist. One way is to present readers with predetermined standards and criteria and some kind of artifact for their application, such as exemplary essays. The other way is to have readers foregather with the aim of discussing the reading at hand, examining sample papers, and arriving at a determination of their own standards and criteria through consensus. The latter procedure was used in the Wisconsin High School Experiment. A form of it is used in at least some of the readings conducted at the Educational Testing Service (ETS). Underlying its use are two assumptions about reader competence: first, if readers are knowledgeable about composition in general, they are specifically competent to arrive at valid standards and criteria; and second, competent people normally are better motivated and achieve better performance when they do things for themselves than when things are imposed upon them.

In spite of which approach is used, readers must have *practice*. The effects of this procedure in raising the level of essay-reader performance was demonstrated a number of years ago by Stalnaker (19).

Practice is indispensable in helping readers work up to the proper *speed* and *rate*. Whether a reader is judging a paper for some specific quality or for general merit, the principle of rate-of-judging is this: up to the point where the prose becomes ununderstandable, the faster the rate and speed the more valid and reliable the judgment. If a reader is competent and if he has been well trained and oriented, his instantaneous judgment is likely to be a genuine response to the thing for which he is looking. But if he is given the time to deliberate, he is likely to accommodate his judgment to tangential or irrelevant qualities which will introduce bias into the judgment. Literary criticism is based upon deliberation, but the judging of a student essay of several hundred words is based upon reflexive decisiveness. This was the strategy used in the Wisconsin High School Experiment. Using a 4-point scale, the readers judged the essays, which averaged about four hundred words in length, at the rate of one essay per minute. An estimate of the effectiveness of this procedure lies in the coefficients of interrater agreement obtained from an analysis of variance of the ratings. These reliability indices for the four topics (284 essays per topic) are:

Topic A	.835
Topic B	.845
Topic C	.801
Topic D	.783

There were four reader judgments per essay, one rating by a judge of each of the four judge types.

Finally, some mention should be made of the factor of reader *fatigue*. At what point it begins to be felt is simply not known. There is reason to believe that it becomes a factor over a number of days, however. From an analysis of the data of a 5-day reading session at ETS, Coffman and others report that the inter-rater agreement was significantly lower on the fifth day than it was on any of the other days (5).

Criteria

A decision must always be made whether to use one general rating (sometimes referred to as the "holistic" method) or a rating of each of a number of specific qualities (the "analytical" method). Coffman's chapter contains a good discussion of the problem, although it is in reference to all kinds of essay tests (4:48–61). Writing tests have certain attributes which make the selection of criteria more ticklish than for other kinds of tests.

There is no question that a composition of adequate length for measurement purposes contains many identifiable specific qualities. Furthermore, everybody would like to see ratings of all qualities contribute to the evaluation of an essay, but there are serious difficulties. These are exemplified in Diederich's study, *Factors in the Judgment of Writing Ability*. The fifty-three readers were invited to annotate each of the three hundred essays they judged. These annotations were classified and together with the ratings were factor-analyzed. What emerged were five factors, designated and described as follows (7:i):

I. Ideas: *relevance, clarity, quantity, development, persuasiveness*
II. Form: *organization and analysis*
III. Flavor: *style, interest, sincerity*
IV. Mechanics: *specific errors in punctuation, grammar, etc.*
V. Wording: *choice and arrangement of words*

Among English teachers these findings have had much influence. They are felt to have some definitiveness because their existence is viewed as having come about "inductively" (the readers were not instructed and were asked to criticize freely). What happens when presumably competent readers judge essays for each of these five specific qualities in a well-designed rating? The answer can be found in the report of a rating performed as part of Page's "Project Essay Grade" (a project investigating the use of computers in analyzing essays, which will be discussed in terms of its main purpose). In this rating, English teachers judged 256 essays for each of the qualities proposed by Diederich, except that the term *flavor* was replaced with *creativity*. For each of the five qualities each essay received eight ratings. The array of intercorrelations of the ratings is reported in Table 1 (15:93).

Table 1.—Page Study Intercorrelations of Diederich Factor Judgments

Trait	1	2	3	4	5
Ideas		.86	.86	.68	.89
Organization	.86		.82	.69	.82
Style	.86	.82		.83	.86
Mechanics	.68	.69	.83		.65
Creativity	.89	.82	.86	.65	

The main problem in the use of two or more specific qualities in essay judging is adequately reflected here. In the words of Page:

> *A constant danger in multi-trait ratings is that they may reflect little more than some general halo effect, and that the presumed differential traits will really not be meaningful. . . . This halo is demonstrated in Table V-9* [*above*], *which shows the intercorrelations among the judged traits of the essays, as rated by eight teachers for each essay. . . . We find a very large halo, or tendency for ratings to agree with each other* (15: 92).

In other words, what happened was that when a judge thought he was judging some specific quality he was actually responding to his general impression or to the overall quality of the essay.

It is interesting that in spite of this evidence Page does not reject entirely the independence of the Diederich factors. An analysis of variance of the ratings was performed, and reported in Table 2 (15:95).

Table 2.—ANOVA Table, Page Ratings with Diederich Traits

Source	SS	df	MS	F
Between Judgments	8,230.305	2,047		
Between Essays	3,791.293	255	14.868	6.002
Error Between	4,439.012	1,792	2.477	
Within Judgments	3,564.414	8,192		
Between Traits	84.212	4	21.053	56.412
Trait x Essay	805.089	1,020	.789	2.115
Error Within	2,675.113	7,168	.373	
TOTAL	11,794.719	10,239		

Page feels that there is importance in "*the significant trait-by-essay interaction, which demonstrates . . . that indeed there is some 'validity' in the different ratings*" (15:94). It should be pointed out that be that as it may, the error term for this particular effect was computed from 7,168 degrees of freedom, which is a whopping number that represents a considerable risk of Type I error (faulty rejection of the null hypothesis).

Scales

We have little knowledge of the effects of the number of points or intervals of writing-test essay-rating scales. One of the few controlled studies on scales was done by this author as a pilot reading in the Wisconsin High School Experiment. The readers were experienced teaching assistants in the Department of English of the University of Wisconsin. The aim of the study was to ascertain whether the distributions of ratings were different and whether inter-rater agreement was different for the 4-point scale and the 6-point scale. There were no such differences, but a time-study revealed that judges were significantly slower with the 6-point than with the 4-point scale (13). This finding was unexpected but not surprising. However, Coffman mentions findings which indicate that in some rating situations judges are as fast with extended scales as they are with contracted ones (4:55).

There are conflicting views about the optimum number of points. One belief is that extended scales (9 or so points or more) allow judges to make finer discriminations, and there is also the belief that contracted scales (3 to 6 points) are more conducive to reliability. Another moot question is the odd versus even number of points, although it obviously is relevant only to contracted scales. It is tempting to think that with a scale with an odd number of points too many ratings will be given at the convenient midpoint, and that a scale with an even number of points will spread the variance too much. Here, as with so many important aspects of essay-rating, we are in need of trustworthy experimental findings.

Topics

Everyone acquainted with essays and with human beings knows that the performance of writers varies from topic to topic. The writer factor has already been discussed. It is presumed that much of this variability can be ascribed to differences between topics.

There is a common opinion that a single topic is inadequate. French states that:

> *The composition test is almost like a one-item test. Some students may happen to enjoy the topic assigned, while others may find it difficult and unstimulating; this results in error* (8:4).

Braddock estimates that in a writing test a student should write at least twice, but in a note he observes that Diederich stated that "*two themes are 'totally inadequate'*" (2:7n). This particular view raises the question whether a test essay is a sampling of writer ability, of topics, or of both. If the first assumption is made, then the topics are of no concern; but obviously there is topic sampling in the construction of a writing test. The concept of validity is critical here because, as can be seen, if a topic is absolutely valid then a sample of one topic is sufficient provided that other sources of error and variability are removed or controlled. The matter of topic validity is perhaps more essential than any other aspect of essay testing in relation to research, since it would seem that the matter is somehow amenable to experimental investigation, and so far there has been none conducted.

Because examiners have long believed that freedom of response lowers the validity and reliability of test essays, the writing test topic is traditionally quite highly structured. The principle which underlies this approach to topics seems to be that if all the writers are given something to say, the effects of knowledge as such will be held more constant than otherwise from writer to writer, and genuine variability due to differences in writing ability will more clearly emerge. The most sophisticated example of this kind of structuring of a topic can perhaps be found in the *Taxonomy of Educational Objectives, Cognitive Domain*, under the heading "Synthesis . . . Production of a Unique Communication" (1:177). This test topic includes a number of short passages (which are not quoted in the *Taxonomy*) dealing with the subject, the "future of private property in America," and 700-odd words of directions which include detailed suggestions of material for exemplification, etc. A simpler example of the same kind of topic is quoted from a report on a college freshman composition course study:

> *TOPIC*
>
> Conventional *is a word frequently used to refer to customary attitudes, beliefs, or actions. In the United States it is a convention for men to be clean-shaven, women to wear a certain amount of make-up, boys to be interested in sports, and girls to be interested in becoming wives and mothers. A person who is* unconventional *in some way departs from the conventions of action or belief of the society of which he is a part.*
>
> *With this explanation in mind, discuss the following statement: 'Convention is society's safeguard, but also its potential executioner.' To what extent and in what ways do you agree with this statement? Use examples and details from*

> *your knowledge and experience to support your conclusions* (9:163).

Of course conventionality is a complicated phenomenon and it would be treated in varying degrees of breadth and profundity, where the knowledge differential between students would be confounding, if the topic were merely something such as, "Discuss conventionality."

Thus, to provide students with something to say in a writing-test topic is to filter out to some extent the factor of subject-matter mastery, which is a different testing objective than is pure writing ability. On the other hand, it seems clear to this author that in giving all students a basic content, and asking them in effect to make something of it, the examination task becomes really a task in logic more than one in pure writing. To take the example above, the topic says to the student, you have been given an explanation of conventionality, you have certain ideas and attitudes toward conventionality based on your knowledge and experience, and you have been given a metaphorical maxim about conventionality, now put these elements together as discourse containing a proposition and proof.

The view taken here is that for a writing-test topic to be valid, it should have the property of filtering out not only differences ascribable to knowledge, but also those arising from fluency in logical operations. The kind of topic proposed as an alternative to the topic highly structured in content, in which all students are given something to say, is a topic in which, actually, students are deprived of something to say. The following two examples are of the latter type:

> *1. Consider these contradictory proverbs: "Look before you leap," and "He who hesitates is lost." Decide which of the two offers better advice, then write an essay in which you defend your choice.*
>
> *2. You have heard the saying, "The best things in life are free." Decide whether this is true or false, then write an essay in which you defend your opinion.*

It is impossible for a writer to deal directly with either of these two topics and to be at the same time intelligent without saying in effect that the idea is so trite and meaningless as to be incapable of any kind of treatment in discourse.

In the Wisconsin High School Experiment the author used four such topics, two of which are the above examples. There is no evidence that they were any more valid than the traditional sort of topic might have been. Again, the conclusion must be voiced that experimentation is sorely needed.

Appearance

The most tangible sources of unreliability and invalidity in essay tests are in the appearance factor, which is quite mechanical but nonetheless critical. Many readers for many years have intuitively been certain of the bias which accrues from a paper's appearance, chiefly its handwriting. Several studies have verified this universal reader impression.

In 1959 an Italian researcher reported the findings of a factor-analytic study of essay dimensionality based on a large-scale essay-rating. Ten judges scored each of 230 high school student essays on seventeen bipolar scales. The first general factor to emerge from the factor-analysis is labeled "Graphic Representation," and is comprised of the specific factors of readability (handwriting), appearance, and esthetic arrangement (18). Luckily this report contains the intercorrelation matrix, which in 1965 this author and Robert Remstad subjected to another analysis, with the results listed in Table 3 (14).

Although it is true that appearance, readability, and esthetic arrangement were among the bipolar scales used by the raters and were thus at the outset of the experiment postulated as comprising a major dimension, it should be noted that their emergence in the analysis is strong and intact.

Other recent studies bear out the importance of the handwriting factor. In an experiment performed on the ratings of law students' essays by a number of professors of law at different universities, the investigators found a correlation of .25 (reported as being significant beyond the .05 level) of the mean rating of each paper by the professors and that of eight handwriting experts (10). In another experiment, where the quality of handwriting was one of several main effects investigated, the F-ratio for the handwriting factor was significant beyond the .01 level. Also, there was a significant interaction between quality of handwriting and order of reading, which probably indicates a tendency for the readers to progressively develop negative bias toward poor handwriting (3). In yet another experiment investigating the handwriting factor, four factor-levels were used: neat handwriting, fair handwriting, poor handwriting, and typewriting. The handwriting effect was significant beyond the .05 level. The ranking of the means for the factor-levels is perhaps more interesting: (1) neat, (2) fair, (3) typed, (4) poor (11). It seems probable that the number of superior ratings given the handwritten papers over typewritten papers reflects a positive bias by way of reward for neatness and legibility.

Without doubt the appearance of each paper in a set is a systematic factor in judging, and there is a tendency for readers to agree in their reactions to appearance. This situation is quite ironic: where the papers that are read are in handwriting, a reliability estimate in the form of a coefficient of inter-rater agreement will be spuriously high because of the correlation of the appearance

Table 3.—First Four Factors Extracted from Analyses of Remondino Data

	Factor Analysis	Incomplete Components	Image Analysis
FACTOR I:			
Originality	.88	.87	.79
Imagination	.85	.87	.80
Maturity	.67	.72	.65
Richness of Ideas	.60	.77	.64
Style	.41	.46	.43
Comprehensiveness	.37	.35	.41
Adequacy of Thoughts	.35	.50	.43
Organization of Ideas		.42	.35
Diction		.36	.35
Accuracy of Facts		.38	
FACTOR II:			
Appearance	.91	.92	.86
Esthetic Arrangement	.91	.94	.87
Readability	.87	.89	.82
FACTOR III:			
Word Forms	.85	.87	.76
Spelling	.78	.82	.68
Syntax	.72	.84	.76
Diction		.64	
Style		.62	
FACTOR IV:			
Adequacy of Thoughts	.68	.67	.69
Richness of Ideas	.63	.37	.39
Accuracy of Facts	.39	.68	.57
Organization of Ideas	.35	.58	.47
Comprehensiveness		.75	.37
Conciseness		.75	
Diction		.40	
Maturity		.37	
Style		.36	

factor in the data. At the same time, the validity is actually lowered, because handwriting ability and writing ability are not the same thing.

The only cure for this condition is to have examination essays typed or put into some other standard printed format. The process is time-consuming and expensive. This author used typewritten essays in the Wisconsin High School Experiment—about 1200 averaging four hundred words each, and the cost per essay was over fifty cents. This may be an appalling prospect to many researchers and to others interested in essay-judging; on the other hand, the case against handwriting seems quite clear.

Design and Methodology

The aim of this aspect of essay testing is to increase the precision of the instrument through controlling both chance factors and systematic factors: to randomize the effects of the former and either to remove or to spread equally among all factors the effects of the latter.

As for examples of these phenomena, the effects of names are systematic, so that names should be removed and replaced with numbers. The effects of script are systematic, but they cannot be removed, they can only be rendered equal for all essays—the usual method being to have the essays typed. The effects of reading-order will be systematic if each reader reads the essays in the same order. Order is transformed into a chance factor through the process of randomizing the order for each reader.

For logical and ultimately for statistical reasons, a normal and independent distribution of the data is desired. The aforementioned procedures help to achieve this condition, but furthermore the readers must work independently. Since they should be reading simultaneously, they should be provided with solitude in some fashion. Also, a reader's score on any essay or any quality should not depend on his score on any other. Thus, forced-choice ratings, in which a reader is obligated to assign a certain ratio of each of the points on a scale, affect the independence of the data in ways which may be vitiating to the analysis (20).

If a writing-test judging is part of an experiment, the design must be especially rigorous. The pretest/posttest design, which in and of itself has serious flaws, seems to be popular in small-scale experimental evaluations of writing performance and writing progress. Perhaps the easiest mistake here is in having the pre- and post-essays rated separately. All essays must be rated at the same time, i.e., as part of the same intact rating session, for, as is obvious, the history of the raters between readings introduces variables which contaminate the data. An instance of the occurrence of this flaw is in a study reported as "Measuring Writing Improvement During an NDEA English

Institute" (21). The duration of the "treatment" (a composition course) was 6 weeks. Two of the three readers judged the pre-essays and the post-essays separately. The analysis (of an unspecified type) of the data turned up a significant difference between the mean rating of the pre-essays and the higher mean rating of the post-essays. The investigator reached the conclusion that "reliable gains were obtained," but in view of the kind of design employed, the inference is unwarranted.

All in all, it seems safe to say that no measurement of essay tests ought to be undertaken until it is certain that the design and the planned procedures are proper, both in terms of the aims of the judging and in the kind of analysis to be performed on the ratings.

Administration

There are two important aspects here over and above those which are standard in all testing. The first is so basic and essential that it goes almost without saying: any writing test must be *impromptu*. All writers whose essays are to be judged as the same sample or population should take the test simultaneously or, if that is impossible, under conditions which are as nearly alike as they can be.

For the second critical factor, *time*, there are two general considerations. First, there is a limit to the minimum time permissible. A writing test should be productive of an essay, not of mere sentences, and it would seem that at least 40 or 45 minutes are required in order to give most students a chance to write an essay. Secondly, two adjustments should always be made relative to administration time: the topic should be compatible with the time, and readers should be trained to accommodate their standards and criteria to writing time. Research is silent on these matters.

Conclusion

No summary will be attempted here. The foregoing discussions have been brief and in most cases do constitute a summary. It is hoped that these discussions reflect the fact that to measure writing ability properly is expensive, complicated, and difficult. The wrong criteria, the wrong topics, the wrong readers, or the improper training, even the innate appearance of the handwritten papers will to some extent vitiate any effort, no matter how exemplary otherwise. Indeed, it seems not unreasonable to say that unless all the necessary conditions can be observed, a writing test should not be attempted. This may seem to be suppressive, but on the other hand it appears wiser to not use judgments of writing ability than to use those which are neither valid nor reliable.

Postscript: Will Computers Help?

This addendum appears because it was just several years ago when Page's "Project Essay Grade," otherwise known as "Analysis of Essays by Computer" (15), was being highly publicized, and people are still asking the question.

When computers become sophisticated enough to read print they will be of much more service than they now are, but that is a mechanical problem which is rather tangential to what is meant by the question. What people want to know is whether there are computer programs available through which indices of the quality of writing can be so analyzed as to predict writing performance better than judges can predict it. The answer is, not yet.

Using thirty predictor variables which are countable by computer (e.g., number of commas, average word length), Page performed a regression analysis with the criterion variable of judges' ratings. The mult-R was .71 (15, 16, 17). In the opinion of this author, that coefficient is not very high. It is lower than what is often reported in the "validation" of objective-tests as so-called writing tests, when essay-ratings are regressed against the objective measures. The simple correlations (uncorrected for shrinkage or for attenuation) of the same judges' ratings used by Page (which came from the Wisconsin High School Experiment) with several typical predictive measures are:

Grade-point-average in English courses	*.62*
ITED #3 (Correctness and Appropriateness of Expression)	*.68*
ITED #7 (Ability to Interpret Literary Materials)	*.63*

It seems clear that the program reported from "Project Essay Grade" has little value for the reliable predicting of essay-ratings.

On the other hand, the computer program which was used is indeed valid, and that is a significant contribution to the research on judging writing ability. It is valid because the variables identified are those of the compositions themselves. In Page's correlation matrix the largest single coefficient is .48, for standard deviation of word length in letters. If other precise indices of similar magnitude can be isolated, the computer program may well be a godsend and a genuine breakthrough.

FOOTNOTES

1. The research reported herein as the "Wisconsin High School Experiment" was conducted pursuant to a contract with the U. S. Department of Health, Education, and Welfare. All references to this study which are not numbered are to further analysis of the data, not reported in the final report or in any publication.

REFERENCES

1. Bloom, Benjamin, and others (eds.), *Taxonomy of Educational Objectives, Handbook I: Cognitive Domain*, Longmans, Green, and Co., New York, 1956.
2. Braddock, Richard; Lloyd-Jones, Richard; Schoer, Lowell, *Research in Written Composition*, National Council of Teachers of English, Champaign, Illinois, 1963.
3. Chase, Clinton I., "The Impact of Some Obvious Variables on Essay Test Scores," *Journal of Educational Measurement*, 5:315–318, 1966.
4. Coffman, William, "Essay Examinations," to appear in Thorndike, R. L. (ed.), *Educational Measurement*, American Council on Education, quoted from draft by permission of the author.
5. Coffman, William; McConville, C. B.; Myers, A. E., "Simplex Structure in the Grading of Essay Tests," *Educational and Psychological Measurement*, 26:41–54, 1966.
6. Copeland, C. T.; Rideout, H. M., *Freshman English and Theme-Correcting in Harvard College*, Silver, Burdett, and Co., New York, 1901.
7. Diederich, Paul; French, John W.; Carlton, Sydell, *Factors in Judgments of Writing Ability*, Educational Testing Service Research Bulletin, Princeton, New Jersey, 1961.
8. French, John W., "Schools of Thought in Judging Excellence of English Themes," 1961 Proceedings of Invitational Conference on Testing Problems, Educational Testing Service, Princeton, New Jersey, 1962.
9. Jewell, Ross M.; Cowley, John; Rhum, Gordon, *The Effectiveness of College-Level Instruction in Freshman Composition*, Final Report of Office of Education Project 2188, University of Northern Iowa, Cedar Falls, 1969.
10. Klein, Stephen P.; Hart, Frederick M., "Chance and Systematic Factors Affecting Essay Grades," *Journal of Educational Measurement*, 5:197–206, 1968.
11. Marshall, Jon C.; Powers, Jerry M., "Writing Neatness, Composition Errors, and Essay Grades," *Journal of Educational Measurement*, 6:97–101, 1969.
12. McColly, William; Remstad, Robert, *Comparative Effectiveness of Composition Skills Learning Activities*, Final Report, U. S. Office of Education Cooperative Research Project 1528, The University of Wisconsin, Madison, 1963.
13. McColly, William; Remstad, Robert, "Composition Rating Scales for General Merit: An Experimental Evaluation," *Journal of Educational Research*, 59:55-56, 1965.
14. McColly, William; Remstad, Robert, *The Dimensions of Composition Annotation*, Final Report, U. S. Office of Education Cooperative Research Project S-216, State University of New York College at Oswego, Oswego, New York, 1965.
15. Page, Ellis B., *The Analysis of Essays by Computer*, Final Report, U. S. Office of Education Project 6–1318, The University of Connecticut, Storrs, 1968.
16. Page, Ellis B., "The Grading of Essays by Computer," Proceedings, 1966 Invitational Conference on Testing Problems, Educational Testing Service, Princeton, New Jersey, 1966.
17. Page, Ellis B., "The Imminence of Grading Essays by Computer," *Phi Delta Kappan*, 28:238–243, 1966.
18. Remondino, C., "A Factorial Analysis of the Evaluation of Scholastic Compositions in the Mother Tongue," *British Journal of Educational Psychology*, 30:242–251, 1959.
19. Stalnaker, John, "The Construction and Results of a Twelve-Hour Test in English Composition," *School and Society*, 39:218–224, 1934.
20. Stanley, Julian C.; Wiley, David, *Development and Analysis of Experimental Designs for Ratings*, Final Report, U. S. Office of Education Cooperative Research Project 789, The University of Wisconsin, Madison, 1962.
21. Veal, Ramon, "Measuring Writing Improvement During an NDEA English Institute," *Journal of Educational Measurement*, 3:303–308, 1966.

New Books from the Publishers

College

Clary, William W. *The Claremont Colleges: A History of the Development of the Claremont Group Plan* (Claremont, CA: Claremont University Center, 1970), xxii + 314 pp. Cloth.

Evans, John Whitney. *The Study of Religion in the State University: An Historical View* (Washington, D. C.: CARA, 1970), vi + 22 pp. Paper $1.

Folger, John K., Helen S. Astin, and Alan E. Bayer. *Human Resources and Higher Education* (New York: Russell Sage Foundation, 1970), xxxii + 475 pp. Cloth $17.50.

Hardee, Melvene D. *Faculty Advising in Colleges and Universities* (Washington, D. C.: American College Personnel Association, 1970), v + 36 pp. Paper $2.50.

Ritchie, M. A. F. *The College Presidency: Initiation into the Order of the Turtle* (New York: Philosophical Library, Inc., 1970), xii + 179 pp. Cloth $5.50.

Robbins, Jeff. *On Balance and Higher Education* (New York: Philosophical Library, Inc., 1970), 102 pp. Cloth $4.50.

Sedgwick, Adam. *A Discourse on the Studies of the University* (New York: Humanities Press, Inc., 1970), 109 pp. Cloth $5.

Small, Samuel A. *Higher Education in the Age of Science* (North Quincy, MA: The Christopher Publishing House, 1970), 167 pp. Cloth $4.95.

Current Issues

Coons, John E., William H. Clune III, and Stephen D. Sugarman. *Private Wealth and Public Education* (Cambridge, MA: Harvard University Press, 1970), xxvi + 520 pp. Cloth $12.50.

Danforth Foundation and the Ford Foundation. *The School and the Democratic Environment* (New York: Columbia University Press, 1970), ix + 115 pp. Cloth $3.95.

Greer, Colin. *Cobweb Attitudes: Essays on Educational and Cultural Mythology* (New York: Teachers College Press, 1970), ix + 70 pp. Paper $3.50.

Kroll, Arthur M. (ed.). *Issues in American Education* (New York: Oxford University Press, 1970), 202 pp. Cloth $6. Paper $1.95.

Morgan, Gordon D. *The Ghetto College Student: A Descriptive Essay on College Youth from the Inner City* (Iowa City: The American College Testing Program, 1970), xii + 64 pp. Paper $3.

Steinfield, Melvin. *Cracks in the Melting Pot: Racism and Discrimination in American History* (Beverly Hills, CA: Glencoe Press, 1970), xx + 354 pp. Paper.

Umans, Shelley. *The Management of Education: A Systematic Design for Educational Revolution* (New York: Doubleday and Company, Inc., 1970), xi + 226 pp. Cloth $5.95.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Cheating Among College Graduate Students[1]

CHARLES H. ZASTROW, University of Wisconsin, Madison

ABSTRACT

Forty-five graduate students received three quizzes in one of their courses, and also filled out the Minnesota Multiphasic Personality Inventory (MMPI) and a questionnaire on cheating. The main findings were: (a) the incidence of cheating among these graduate subjects was found to be at least as extensive as reported in prior studies among undergraduate students—at least eighteen of the forty-five (40%) graduate subjects cheated on the quizzes; (b) pressure to obtain good grades was the main reason subjects felt they cheated in the past; (c) there was a lack of consensus whether certain behaviors constitute cheating; (d) no significant personality differences were found between cheaters and non-cheaters; (e) the MMPI results provided evidence in support of the doctrine of specificity of moral behavior.

ALTHOUGH numerous studies have investigated cheating among elementary, secondary, and undergraduate college students, a review of the literature did not reveal any studies conducted on graduate students. The present study is designed to investigate the following aspects of cheating among graduate students in order that cheating may be better understood at all educational levels:

(1) The incidence of cheating.
(2) The reasons for cheating.
(3) The kinds of behaviors considered to constitute cheating.
(4) Differences in personality characteristics between cheaters and non-cheaters.

The reported incidence of cheating among undergradute college students ranges from 24 to 66 percent in different investigations (2, 3, 4, 6, 8, 10, 13). Studies of elementary, secondary, and undergraduate college students also reveal that cheating becomes more prevalent in the higher grades. James (9) in a study using the same tests of honesty at different educational levels, found an increase in cheating from the third grade (12%) to the ninth grade (93%). Steiner (11) found ninth-grade students were more dishonest than eighth-grade students at the same school. Parr (10) found an increase in cheating with the age of the student from the 17-year level to the 21-year level. The trend for cheating to become more prevalent in the higher grades suggests cheating *may* be widespread among graduate students.

However, several studies of undergraduate college students indicate cheating is less prevalent among students with high grades (2, 4, 5, 8, 10). Drake (4), for instance, in a study of 126 college females found no "A" students cheated, while 4 percent of the "B" students, 23 percent of the "C" students, 75 percent of the "D" students, and 67 percent of the "F" students cheated. Since most graduate students are A or B students, this trend would suggest cheating *may not* be widespread among them.

Investigations of the reasons for cheating have indicated cheating is associated with a difficult test (8); lack of supervision, pressure to obtain good grades, a poorly constructed test, a poorly organized course (12); and academic handicaps, such as being involved in numerous extracurricular activities or having a part-time job (10).

Previous studies also indicate there are differences in personality characteristics between cheaters and non-cheaters. Strang (12) found cheaters to be more neurotic, introverted, dominant, and less self-sufficient than non-cheaters. Hetherington (6) found cheaters use more repressive defenses than non-cheaters, and manifest a passive-dependent mode of adjustment.

Method and Procedure

Forty-five of the entering first-year social work graduate students enrolled in one of the required courses at a large Midwest university unwittingly served as Ss in the study. During the semester they received three unannounced quizzes. Each quiz was two pages in length, with the pages stapled together. On the first page were true-false questions, with the answers having to be recorded in a designated space to the left of each question. On the second page was an essay question. After completing the quizzes in class, Ss were instructed to separate the two pages and hand in the second page in order to allow the instructor to grade the essay question. (The second page was never graded, and the quizzes, unknown to the subjects during the research project, never influenced their final grade in the course.) The Ss were then informed to take the first page home; grade it (indicating which true-false questions they answered incorrectly); and return it to the instructor within a 6-day period. The correct answers were relatively easy to locate since the questions were taken directly from the required textbook for the course.

Unknown to the Ss, the back of the first page of each quiz contained a white, invisible coating of powdered lead carbonate. When the students answered the true-false questions on the first page, an invisible impression from the white powder was transferred to the second page. Immersing the second page into an atmosphere of hydrogen sulfide gas caused the white powder to turn dark brown; thereby revealing a visible copy of what was recorded on the first page. This copy was as plain as a carbon copy. With this procedure it was possible to have an exact copy of the Ss' answers to the true-false questions before they graded their own papers. By comparing the Ss' answers before they graded their own papers to the answers on the first page after they graded their papers, it was possible to detect those Ss who changed some of their answers while grading their papers. Cheating was defined to have occurred when a S changed an answer while grading his paper. This particular method for detecting cheating was adapted from procedures originally developed by Howells (7).

The experimenter being a second-year graduate student in the same master's program as the Ss, had numerous informal conversations with them which revealed they were not suspicious of the rather strange test-taking procedure of being allowed to grade their own papers; the procedure was generally interpreted as being a technique to help them learn the material in the textbook.

After the Ss graded and returned the last of the three quizzes, the instructor informed them of their participation in a study on cheating, and also told them of the particular method used to detect cheating. A questionnaire was then administered which: (1) asked the students if they had ever cheated on tests during the years they attended school; (2) if they indicated they had cheated, asked them to list the reasons for cheating; and (3) asked questions regarding the behaviors they considered to constitute cheating. Ordinarily, students would be expected to be defensive about filling out such a questionnaire for fear of being labeled a "cheater." However, since the questionnaire was administered after the Ss were informed of their participation in a study on cheating, it was anticipated they would be fairly honest in filling out the questionnaire since they knew the quizzes, rather than the questionnaire, were being used to identify those who cheated.

In an attempt to discover if there are differences in personality characteristics between cheaters and non-cheaters, the group form of the MMPI was administered to the Ss. The inventories were scored on eleven standard scales (validity, lie, hypochondriasis, depression, hysteria, psychopathic deviate, paranoia, psychasthenia, schizophrenia, hypomania, and social).

Results and Discussion

Incidence of Cheating. One of the rather striking findings is the large percentage of Ss detected to have cheated on at least one of the three quizzes: at least eighteen of the forty-five (40%) cheated. The exact number of students who cheated is unknown because it could not be ascertained for several reasons whether fourteen of the Ss cheated. In order to be classified as a non-cheater, the design of the study required that a student did not change any of his answers while grading all three of the quizzes. Six of the students were placed in an "uncertain" category because they were absent when one of the three quizzes was administered, even though they took and did not cheat on the other two quizzes. Four of the Ss separated the two pages of a quiz before filling out the first page in class, with the result being no impression of how they answered the true-false questions was made on the second page. Furthermore, in three cases, the impression left on the second page was too faint to distinguish how Ss had answered the true-false questions in class. And, in one case, a S failed to return a quiz after taking it home to grade.

Even though the exact percentage of Ss who cheated is unknown, the results show at least 40 percent cheated. This is a sizeable percentage, and suggests that the incidence of cheating among graduate students is at least as extensive as among undergraduate students where the reported incidence of cheating ranges from 24 to 66 percent in different investigations (2, 3, 4, 6, 8, 10, 13).

One of the questions on the questionnaire was closely related to the incidence of cheating, as it asked Ss whether they had ever cheated on a test during the years they attended school. Forty-two of the forty-five respondents indicated they cheated in the past. Furthermore, the three respondents who said they did not cheat were detected to have cheated on the quizzes. Therefore, all forty-five subjects have cheated sometime during their school years.

Reasons for Cheating. Prior studies of the reasons for cheating have indicated pressure to obtain good grades; a poor test; a poorly organized course (12); a difficult test (8); and handicaps, such as having a part-time job or being involved in numerous extracurricular activities (10) may encourage cheating.

The present study confirms most of these reasons and adds several others. The subjects listed on the questionnaire, the reasons listed in Table 1 for cheating. Of the forty-five subjects, seventeen listed "pressure to obtain good grades," which was by far the most frequently mentioned reason. "Being unprepared for a test," listed by seven students, was second.

The above reasons indicate the high value students ascribe to receiving good grades. Good grades are not only recognized as being a necessity for academic advancement and eventual job opportunities, but also appear to have psychological meaning to students; such reasons as "fear of self-devaluation in competition" and "lack of confidence in own abilities" suggest that students judge, to some extent, their abilities (i.e., their worth) by the grades received in competition with other students.

Some of the other reasons ("poor instructors," "ambiguous and unfair tests," "inadequate time to complete test," and "unannounced tests") may be rationalizations for cheating.

Forms of Cheating. Many different types of behaviors are classified as cheating. Using crib notes, looking at another's paper while taking an exam, and changing answers while grading one's own exam are widely recognized as cheating. But there is a lack of consensus over whether certain other behaviors constitute cheating. The subjects' responses to the questionnaire revealed that 95 percent did not consider using materials from previous papers when writing an assigned paper to be cheating; 74 percent did not think reading a condensation of a novel when assigned the full version is cheating; 52 percent did not believe failing to appear for an exam at the scheduled time is cheating; and 23 percent did not consider listing unread primary sources in bibliographies when writing a term paper to be cheating. These results show there is a lack of consensus among graduate students over whether certain behaviors should be considered to be cheating. Difficulties may arise when an instructor views such behaviors as cheating, while some of the students do not.

Table 1.—College Graduate Students' Reasons for Cheating

Reason	Number of Subjects Indicating the Reason
Pressure to obtain good grades	17
Being unprepared for a test	7
Opportunities to cheat	5
Desire to impress teachers and/or parents	5
Desire to impress peers	4
Fear of self-devaluation in competition	4
Unannounced tests	4
Ambiguous and unfair tests	3
Difficult tests	3
Poor instructors	2
Lack of confidence in own abilities	2
Lack of supervision	2
Peer pressure to cheat	2
Inadequate time to complete test	1
Unstimulating course materials	1

Personality Characteristics of Cheaters. Scores obtained on eleven standard scales of the MMPI were compared, by t-tests, to the scores obtained by non-cheaters. No significant differences were found. These results are somewhat surprising since investigations of cheating among elementary and secondary school children, and among undergraduate college students have found differences in personality characteristics between cheaters and non-cheaters. (6, 12).

The MMPI results also have implications for the specificity–generality of moral behavior controversy (1). The generality of moral behavior formulation hypothesizes the existence of a general trait of honesty. This view predicts that a person will be consistent in his behavior over many different kinds of situations; i.e., it predicts a person who cheats in one situation is not only likely to cheat in other situations, but is also highly likely to lie, not feel guilty, steal, embezzle, etc. If a general trait of honesty exists, students who were detected to have cheated should also receive higher scores on the Lie (L) scale and on the Psychopathic Deviate (Pd) scale. (They would be expected to receive a higher score on the Pd scale because one of the main characteristics of psychopaths is they experience little conscious guilt.) The t-test scores, however, reveal no significant differences between cheaters and non-cheaters on these two scales, and, therefore, do not support the existence of a general trait of honesty.

Instead, these MMPI results support the opposite point of view; "the doctrine of specificity of moral behavior." This view theorizes that a person acts in each situation according to the way he has learned to act in similar situations in the past. The predictability of one's behavior from one situation to another depends, not on the notion of a general trait of honesty, but on the number of identical elements which the two situations share. This formulation does not accept the concept of "honesty" as being a valid character trait, but instead maintains there are many different kinds of specific behaviors which tend to be independent even though they are grouped together under the same concept. Thus, this view holds that knowing whether a person has cheated on tests will not permit one to predict how that person will act when given an opportunity to be dishonest in a different situation.

If a general trait of honesty does not exist, as the findings in this study suggest, there is no scientific basis for the common sense prediction that graduate students who cheat on tests are more likely than non-cheaters to be dishonest in either their specialized area of academic training, or in any other area.

FOOTNOTE

1. This research was orginally done in partial fulfillment of the requirements for the master's degree in social work. The author wishes to thank his advisor, Morton Perlmutter, for his valuable criticisms and suggestions in designing the study and for his generous time and assistance in conducting the project.

REFERENCES

1. Burton, R. V., "Generality of Honesty Reconsidered," *Psychological Review*, 70:481–499, 1963.
2. Campbell, W. G., "Measurement in Determining the Personality and Behavior of the College Cribber," *Education*, 53:403–408, 1933.
3. Campbell, W. G., *A Comparative Investigation of the Behavior of Students Under an Honor System and a Procter System in the Same University*, University of Southern California Press, 1935.
4. Drake, D. A., "Why Students Cheat," *Journal of Higher Education*, 12:418–420, 1941.
5. Hartshorne, H.; May, M. A., *Studies in Deceit*, Macmillan, New York, 1928.
6. Hetherington, E. M.; Feldman, S. E., "College Cheating as a Function of Subject and Situational Variables," *Journal of Educational Psychology*, 55:212–218, 1964.
7. Howells, T. H., "A Method of Secretly Obtaining Duplication of Writing," *Journal of Abnormal and Social Psychology*, 27:334–335, 1932.
8. Howells, T. H., "Factors Influencing Honesty," *Journal of Social Psychology*, 9:97–102, 1938.
9. James, H. W., "Honesty as a Character Trait Among Young People," *Journal of Educational Research*, 26: 572–578, 1933.
10. Parr, E. W., "The Problem of Student Honesty," *Journal of Higher Education*, 7:318–326, 1936.
11. Steiner, M. A., "Cheating in School," *School Review*, 40:535–546, 1932.
12. Strang, R., *Behavior and Background of Students in College and Secondary Schools*, Harper, New York, 1937 pp. 75–81.
13. Yepsen, L. N., "Reliability of Self-scored Measures," *School and Society*, 29:681–682, 1929.

BOOK REVIEWS

ROBERT E. CLASEN
Book Review Editor

What is Philosphy of Education?

Lucas, Christopher J., Macmillan Company, New York, 1969, 313 pp.

BASICALLY, THIS IS AN anthology of supplementary readings for students in philosophy of education, it is one of the most intriguing books I have ever read and also one of the most disappointing. The book talks all about philosophy of education, but when I look for defended answers to vital questions, I find the writer more skittish than a politician. In fact, the prevailing idea is that there is no right philosophy and no commonality in essentials among leading philosophies. This tempts educators to conclude that there is no use in wasting their time with philosophy as philosophers appear to have done! To illustrate, I list five fundamental questions, briefly state my tentative answers, and look for answers in this anthology:

Question 1. What is a person's ultimate goal? While contemporary religions stress Nirvana, Heaven, and Hell, my personal guess is extinction except as one lives in his progeny or continuing influence. The anthology gives no answer to a question which has been a fundamental deciding point in the educational psychology of teachers the world over and for all times.

Question 2. What should be a person's goal during his life? The world's religions suggest that we so live as to conform to Deity's wishes as revealed to Buddha, Jesus, etc. Many in the rising generation are seriously questioning some of these prescriptions, claiming that they were strained through the whiskers of the elders in whom the desires of youth have faded. The anthology by Lucas provides some absorbing discussion of the purposes of philosophy but no real attempt to answer this kind of question.

Question 3. Do self-urges disregard the urges of others? My guess is that enlightened self-urges take into intimate account the urges of others. Each person has the self-urge to avoid pain from the kick-back when the urges of others are violated. Each person also has the urge to secure the approval of others. He may even have the urge to get satisfaction from satisfying the urges of others. The anthology by Lucas provides no consideration of this question except for a brilliant discussion of the urgent need for philosophy to find acceptable methods even as science and mathematics have developed methods yielding a rich harvest.

Question 4. In choosing a course of action, what percent of weight should a person give to his own urges and what percent to the urges of others? I think nature was wise in providing that each person act 100 percent for himself. He can judge rather accurately what will satisfy himself and what will be the effect of kickback, approval or disapproval. He will be less accurate in weighing the complex and often concealed urges of others. The anthology by Lucas provides no specific answer to the question except by distant inference. It does contain an excellent presentation of the educational philosophies of Dewey, Kilpatrick, Counts, Royce, Whitehead and others. These philosophies, as presented, do not provide a specific answer to the foregoing question which is of great concern to educators.

Question 5. Which should be given priority, average or total satisfaction? I lean toward average satisfaction but I am still waiting for help in reaching a firmer conclusion. Parents with two children can provide them with an average satisfaction in life of, say, 80 points on a scale of 100; their total is then 160. These same parents might have twenty children. Their average would probably drop to, say, 60 points, but the total is 1,200. Should the parents aim at a high average or a high total? Because of the present population explosion this matter of average or total is of acute concern not only to parents but also to statesmen, for the present population can be increased many times without lowering the total units. Again Lucas provides no consideration of this educationally-relevant philosophic question.

Surely educational philosophers must have debated the five questions. Probably Lucas chose selections less to present answers to questions than to create among educators a philosophic attitude of mind, an awareness of the major philosophic systems and how they relate to education. This was well worth doing and was well-done.

William A. McCall
Emeritus Professor of Education
Teachers College
Columbia University

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

A Longitudinal Study Examining Prediction of Doctoral Success: Grade Point Average as Criterion, or Graduation vs. Non-Graduation as Criterion

JOHN D. WILLIAMS, STEVEN D. HARLOW, and DEL GAB
University of North Dakota, Grand Forks

ABSTRACT

A comparison of two different criteria of doctoral success is made using the criterion of doctoral grade point average (GPA) and the criterion of whether or not the student received a doctoral degree. The study is longitudinal in nature, and was restricted to one department (Education) at the University of North Dakota over a 7-year period. The GRE Quantitative correlated .34 with the graduated-not-graduated criterion, and —.01 with the doctoral GPA. Other variables had even more divergent results: the multiple correlation coefficient for the graduated-not-graduated criterion was .650, and the multiple correlation coefficient for the doctoral GPA criterion was .506. Hence, it was concluded that the two measures of doctoral success were in fact different.

PREDICTING success in a doctoral program would be a most worthwhile accomplishment, providing it could be done with a reasonable degree of accuracy. One problem that has plagued researchers in this important area has been defining the criterion of success. For example, first semester GPA (doctoral) or cumulative doctoral GPA have often served as the criterion. Typically, this type of criterion has proved evasive to predictability, with correlations (or multiple correlations) generally falling below .50. Part of the reason for this lack of predictability is in the relatively short range of grades; quite often, graduate courses are restricted to grades of A's and B's with only a few C's.

Setting aside the concern for the restricted grade range for the moment, it would seem necessary to ask if the criterion is meaningful. If the *raison d'être* of a doctoral program is to be concerned with graduating doctoral students, then a much more important criterion becomes viable: whether or not a student will graduate. This rather simple, yet very important criterion has been praised for its appropriateness (1, 3), but has not really been fully investigated. Several reasons for this neglect can be understood. First of all, using the criterion of graduated-not-graduated immediately implies studies of a longitudinal nature. Perhaps this reason alone has given precedence to the use of the accessible doctoral GPA. Also, investigators who have tried using this approach have not always met with success. Roemer (2), in a 9-year study of predictors of success in medical school, found little relationship, either singly or in combination, among his predictors and medical school GPA. Wright (3) conducted a longitudinal study of considerable depth between 1951–1962. Of the sixty-three variables he investigated, only one variable proved to be of significance. Age was found to be negatively related to the criterion of whether or not the student received a graduate degree.

Several points should be made concerning a doctoral program, particularly in education. First, a dissertation is a usual requirement; thus, a person could conceivably achieve good grades yet never complete a dissertation. Second, a doctoral program is typically conducted for the most part in a single department, rather than being a university-wide activity. Those predictors that may have some validity in a doctoral program in physics, for example, may have very little validity for a doctoral program in education. Thus, it would seem to be worthwhile to examine success within a single department, using the graduated-not-graduated criterion.

Population and Methodology

The cumulative files of all students who had been admitted to the doctoral program in the Education Department between June, 1962, and June, 1967, at the University of North Dakota were examined, using information available as of February, 1969. Three groups of students were defined:

(1) students who had received the doctorate,
(2) students who had not received the doctorate, and who had not been in attendance nor registered for dissertation credit, either during a regular semester or the summer term, for the preceding 21 months,
(3) students who had not received the doctorate, but who had either been in attendance or registered for dissertation credit, either during a regular semester or the summer term, at least once during the preceding 21 months.

The third group was considered to be, for our purposes, continuing toward the objective of receiving a doctorate. The third group was thus not included in the present study.

With this definition and population, the number of graduates was thirty-three, and the number of non-graduates was fifty-one. The following information was secured:

1. Area of concentration (either teacher education = 1, or administration = 0)
2. Years between bachelor's degree and doctoral admission
3. Years between master's degree and doctoral admission
4. Whether or not master's degree was from UND (yes = 1, no = 0)
5. Age at admission
6. Accumulative GPA for bachelor's degree
7. Accumulative GPA for master's degree
8. Number of courses in education before starting doctoral program
9. Number of years of educational experience
10. Graduate Record Examination (GRE) Verbal score
11. GRE Quantitative score
12. GRE Advanced Test in Education score
13. Miller Analogies Test (MAT) score
14. Degree sought (PhD = 1, EdD =0)
15. Sex (male = 1, female = 0)
16. Cumulative doctoral GPA
17. Whether or not the student was graduated

The first fifteen variables served as predictors, and the last two variables served as criteria.

Results and Discussion

In Table 1 are the zero-order correlation coefficients of the fifteen predictor variables and the two different criteria. The multiple correlations for both are also included.

Table 1.—Correlations of Background Variables, Test Data, and Two Different Criteria of Doctoral Success

Variable	Criterion	
	Doctoral GPA	Graduated-Not-Graduated
Area of Concentration	.09	-.04
Years Bet. Bach. Degree and Doctoral Admission	-.05	.09
Years Bet. Master's Degree and Doctoral Admission	-.12	.07
UND Graduate	-.14	.36**
Age at Admission	-.15	.02
GPA for Bachelor's Degree	.13	.02
GPA for Master's Degree	-.01	.22*
No. Courses in Education	.01	.18
No. Years Educ. Experience	.05	.11
GRE - Verbal	-.01	.08
GRE - Quantitative	-.01	.34**
GRE - Advanced Test in Educ.	.08	.34**
MAT	-.03	.10
Degree Sought	.01	.11
Sex	-.14	.22*
Multiple Correlation, All 15 Variables	.506	.650**

*Significant at .05 level
**Significant at .01 level

It becomes apparent, when inspecting Table 1, that the two criteria of doctoral success are definitely *not* the same. The single highest zero-order correlation on the graduated-not-graduated criterion (with graduated = 1, not-graduated = 0) is whether or not the individual graduated with his master's degree from UND.

Other notable discrepancies occur as well. At the master's level, the GPA correlated .22 with the binary criterion of graduated-not-graduated, but correlated —.01 with doctoral GPA. If the reader realizes that the student who received his master's degree from another school does not have his previous grades included in his doctoral GPA, then this result seems more plausible. The test scores (Graduate Record Examination scores, and the Miller Analogies Test scores) also allow interesting interpretations. In every case, they show a higher validity for the graduated-not-graduated criterion than they do for doctoral GPA. Only the GRE-Advanced Test in Education has a positive correlation with doctoral GPA. The largest difference is with the GRE Quantitative with $r = -.01$ with the criterion of doctoral GPA, and $r = .34$ with the binary criterion of graduated-not-graduated.

It is also noted that the criterion of graduated-not-graduated has a higher multiple correlation with the fifteen variables than does the criterion of doctoral GPA. For the graduated-not-graduated criterion, $R = .650$, with $R^2 = .423$; for doctoral GPA, $R = .506$, with $R^2 = .256$. Not only is the criterion of graduated-not-graduated more important in the future of the recipients (or non-recipients) but in this case, it is more predictable.

Table 2 contains the means and standard deviations for the fifteen variables for both groups—those who have been graduated and those who have not been graduated. It gives a more descriptive view of information than is found in Table 1. Whether or not the student had received his master's degree from UND can now be seen to have a rather dramatic effect. Of those who have been graduated, 64 percent received their master's from UND. The differences on the GRE Quantitative and Advanced Test in Education are marked. The mean of the graduates on the GRE Quantitative is 566.05 as compared with the non-graduates' mean of 504.27. It should be remembered that the GRE Quantitative was unrelated to success on the doctoral GPA, and yet seems highly related to success when defined as graduated or not graduated. One interpretation is that quantitative ability is not particularly important in the course work, but the student with weak quantitative ability is less likely to complete a dissertation which is statistical and/or experi-

Table 2.—Means and Standard Deviations on 15 Variables of Students who Have Graduated or Not Graduated

Variable	Graduated N=33 Mean	Graduated N=33 S.D.	Not Graduated N=51 Mean	Not Graduated N=51 S.D.
Area of Concentration	0.73	0.44	0.76	0.42
Years Bet. Bach Degree and Doctoral Admission	9.88	4.81	8.94	4.92
Years Between Masters Degree and Doctoral Admission	4.42	3.37	3.94	3.52
UND Graduate	0.64	0.48	0.28	0.44
Age at Admission	35.06	5.94	34.82	6.86
GPA for Bach. Degree	2.75	0.39	2.73	0.39
GPA for Masters Degree	3.54	0.24	3.43	0.25
No. Courses in Educ.	10.73	5.19	8.92	4.46
No. Years Educ. Experience	9.45	5.56	8.14	6.07
GRE - Verbal	514.84	75.40	502.67	70.95
GRE - Quantitative	556.06	101.44	504.27	71.04
GRE - Advanced Test in Educ.	589.09	58.89	553.82	38.16
MAT	51.73	8.54	49.55	11.40
Degree Sought	0.12	0.32	0.06	0.23
Sex	1.00	0.00	0.88	0.32
Doctoral GPA	3.73	0.13	3.51	0.75

mental in nature, and consequently may not complete his degree.

Considering the GRE Advanced Test in Education, the graduates have a higher mean (589.091 as compared with 553.823), which is in the expected direction.

The Miller Analogies Test (MAT) was basically unrelated to either criterion; however, the graduates have a slightly higher mean.

Tables 3 and 4 contain the steps in the stepwise regression (backward elimination procedure) for the two criteria. If the tables are read from bottom to top, the more important predictors of the two criteria can be ascertained.

The importance of the GRE Tests as predictors of graduation can be seen in Table 4. On the other hand, they are not particularly useful in predicting doctoral GPA. In fact, the GRE-Quantitative is the *first* variable to be eliminated with GPA as the criterion, but it is *next to last* to be eliminated with graduation as the criterion. It would be fair to say that the two criteria are in fact very different, and should not be confused.

The Miller Analogies Test (MAT) is the second variable dropped regardless of the criterion. Evidently, success on the MAT is unrelated to success by either definition with the given population.

Table 5 gives the results of the discriminant analysis, with the two groups being defined on the graduated-not-graduated criterion.

Implications

The solution to the problem of selecting a criterion to measure the success of various predictors with doctoral programs has previously been seen to be elusive. While other criteria might be used, two common criteria are GPA and whether or not the student was graduated. If the latter criterion is at all predictable, it has much appeal in its favor. The present study has shown that if a longitudinal study is limited to a single department, *promising predictability* can be found with the variable of graduated-not-graduated as the criterion. It was found, generally, that the Graduate Record Examinations have the greatest predictive value, although other background variables do increase prediction somewhat. It would seem that the two criteria are definitely different. The Miller Analogies Test has little relationship with either criteria, while elimination of this test might seem a rash judgment, certainly less emphasis in selection with the MAT would seem appropriate.

It should be pointed out that the criterion problem has by no means been solved. Even in defining the criteria for this study, some difficulty of

Table 3.—Stepwise Backward Elimination Procedure with Doctoral GPA as the Criterion N = 84

Step	Variable Eliminated	Multiple Correlation	Significance Level
1	None	.506	$p > .05$
2	GRE-Quantitative	.503	$p > .05$
3	MAT-Test Score	.496	$p > .05$
4	Sex	.485	$p > .05$
5	Number of courses in education before doctoral admission	.471	$p > .05$
6	Years between Bachelor's and Doctoral Admission	.457	$p > .05$
7	Master's degree received at UND	.444	$p < .05$
8	Years between master's and doctoral admission	.420	$p > .05$
9	Accumulative GPA for master's degree	.402	$p > .05$
10	Accumulative GPA for bachelor's degree	.385	$p < .05$
11	GRE-Verbal	.365	$p < .05$
12	GRE-Advanced	.349	$p < .05$
13	Area of concentration Teacher Ed. or Administration	.323	$p < .05$
14	Degree sought Ph.D. or Ed.D.	.257	$p > .05$
15	Number of years of educational exper.	.151	$p > .05$
16	Age	----	----- ---

interpretation might occur. Some of those students who have been considered as unsuccessful (i.e., not enrolled within the last 21 months) may eventually receive a doctorate. Also, with the GPA criterion, there is a likelihood that a positive relationship will exist between number of courses attempted and GPA. This would occur as an artifact if it would occur for no other reason. If a student were eliminated on the basis of his GPA, his number of courses attempted would be less than someone who completed the doctorate. Likewise, students who are not officially eliminated by their GPA may perceive their chances of receiving the doctorate as rather low, and discontinue their doctoral studies. At any rate, the amount of time spent on the doctorate has not been specifically controlled in the present study. There is sufficient evidence, however, to recommend the graduated-not-graduated criterion as opposed to the doctoral GPA as being the criterion of doctoral success.

Table 4.—Stepwise Backward Elimination Procedure with Graduated or Not-Graduated as the Criterion N = 84

Step	Variable Eliminated	Multiple Correlation	Significance Level
1	None	.650	$p < .01$
2	Years between master's and doctoral admission	.650	$p < .01$
3	MAT-Test Score	.650	$p < .01$
4	Accumulative GPA for bachelor's degree	.650	$p < .01$
5	Years between bachelor's and doctoral admission	.648	$p < .01$
6	Age	.646	$p < .01$
7	Area of concentration Teacher Ed. or Administration	.641	$p < .01$
8	Accumulative GPA for master's degree	.634	$p < .01$
9	Sex	.629	$p < .01$
10	Number of years of educational exper.	.618	$p < .01$
11	Number of courses in education before doctoral admission	.591	$p < .01$
12	Master's degree received at UND	.545	$p < .01$
13	Degree sought Ph.D. or Ed.D.		
14	GRE-Verbal	.499	$p < .01$
15	GRE-Quantitative	.435	$p < .01$
16	GRE-Advanced	.342	$p < .01$

Table 5.—Discriminant Analysis of Doctoral Students into Graduated or Not Graduated Groups

		ACTUAL	
		Graduated	Not-Graduated
PREDICTED	Graduated	26	9
	Not-Graduated	7	42

Mahalanobis $D^2 = 66.06645$ $p < .01$

REFERENCES

1. Angoff, William H., "Graduate School Guessing Game," *Science*, 164:1117–18, 1969.
2. Roemer, R. E., "Nine-Year Validity Study of Predictors of Medical School Success," *The Journal of Educational Research*, 59:183–85, 1965.
3. Wright, Charles, "Success or Failure in Earning Graduate Degrees," *Sociology of Education*, 37:73–97, 1964.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

The Relationship Between Ego Identity Status and Satisfaction with College[1]

ALAN S. WATERMAN
Rensselaer Polytechnic Institute

CAROLINE K. WATERMAN[2]
State University of New York at Albany

ABSTRACT

Two studies were conducted to investigate the relationship between Erikson's concept of ego identity and student attitudes toward college. The results of Study 1 indicated that students who were in the process of going through an identity crisis over occupational choice had the least favorable evaluations of their education while students who had never experienced a crisis but were, nevertheless, committed to a vocational goal held the most favorable attitudes. Study 2 demonstrated that students who went through crisis over choice of major field while in college showed more negative attitudes than did students not experiencing a crisis during their college years. These findings suggest that the association of crisis with college is an important factor in student dissatisfaction.

THE DIFFICULTIES colleges and universities have in meeting the needs of their students is reflected both in the rate of student attrition and in the frequency with which students criticize their schools. A review of the research on college dropouts by Summerskill (8) reveals that 40 percent of entering students never receive a degree. Of those students dropping out of college, less than half do so for reasons of academic failure or ill-health. For those students remaining in college until graduation there is evidence of a significant degree of dissatisfaction. Trent and Medsker (9) found that while college seniors generally felt their college experiences were satisfying, almost one-third voiced serious criticism concerning the quality of their college faculty and the administration's treatment of students.

The level of a student's satisfaction with his school can be thought of as a function of two classes of variables: (a) variables of the college and (b) variables of the student. The present studies deal with the latter type of variable. Specifically, they are concerned with the variable of ego identity status.

The theoretical groundwork for the concept of ego identity has been developed by Erikson (3, 4). He has formulated a theory of personality development in which each stage of a person's life is seen as involving some important crisis resulting from a socially-posed problem. The different components of personality are determined by the manner in which each of these crises is resolved. The crisis which is important during adolescence involves the establishing of a stable sense of identity. The adolescent in our society is faced with a decision about what type of person he will become. This includes both making a decision about an occupation and arriving at a stable set of beliefs, for example, about religion and politics. Erikson spoke of a successful resolution of the crisis as the achieving of an ego identity. A failure to successfully resolve the crisis was seen as resulting in identity diffusion.

Marcia (5) elaborated further on Erikson's ideas regarding identity. He proposed a 4-category classification of this construct: (a) *Identity Achievement*, (b) *Moratorium*, (c) *Foreclosure*, and (d) *Identity Diffusion*. According to Marcia, these various ego identity statuses can be categorized on the basis of two components: crisis and commitment. Crisis refers to a period of struggle or active questioning in arriving at a set of beliefs and an occupational choice. Commitment refers to a firm, unwavering choice regarding beliefs and occupation.

An *Identity Achiever* (IA) is someone who has gone through a period of crisis and has developed a stable commitment to an occupation and a belief system. The term *Moratorium* (M) is used to refer to a person who is currently in a state of crisis, and is actively seeking among alternatives in an attempt to arrive at a choice. A person is classified as a *Foreclosure* (F) if he has never experienced a crisis but nevertheless is committed regarding occupation and beliefs. The occupation and beliefs chosen are generally ones which reflect the wishes of his parents or other authority figures. The category of *Identity Diffusion* (ID) includes individuals who are not committed to anything and who are not actively trying to make a commitment. These people may never have been in crisis, or they may have had a period of questioning but were not able to resolve it and subsequently emerged without making a choice.

This concept of ego identity, particularly those aspects concerning occupational choice, would appear to have important implications for the student's evaluation of his college experiences.

Douvan and Withey (2) found a tendency among high school boys to describe their college plans in terms of specific occupational fields. Beardslee and O'Dowd (1), in making a similar observation, noted that few freshmen have a realistic conception of what the work in their chosen fields involves. As a consequence, many students, after arriving at college, are forced to reevaluate their vocational plans; i.e., they enter a period of crisis. That this process of questioning can be stressful to the student is indicated by the finding that M's show higher levels of anxiety than other statuses (6). It would appear plausible to expect that students who go through this stressful occupational crisis while at college would tend to have relatively negative attitudes toward their school, since it has been associated with this crisis. On the other hand, students who are committed to an occupation when they arrive on campus and who do not subsequently experience an occupational crisis will not associate the stress of crisis with college and may perceive the school as helping them fulfill their vocational plans. They should, therefore, be relatively satisfied with their college experience. Thus, it can be hypothesized that F's on occupational choice, will tend to have the most favorable attitudes toward college while M's will tend to have the least favorable attitudes.

STUDY 1

A study testing this hypothesis was conducted at Rensselaer Polytechnic Institute (RPI). Since RPI is a technologically oriented university, the question of vocational choice is particularly important for the student in his decision to come to the school. A large majority of the freshmen anticipate careers in engineering or physical science. The attrition rate is somewhat lower than the national level, averaging approximately 30 percent. However, for each entering class, there occurs a substantial transfer of students from the engineering and science programs to programs in management, humanities, and social sciences. These shifts generally occur during the student's first 3 years on campus.

Method

Subjects. The sample consisted of fifteen freshmen and seventeen juniors, all of whom were taking the engineering program at RPI.

Procedure. All Ss were given a semi-structured interview to assess their ego identity status. The form of the interview was the same as that described by Marcia (5). The topics covered included the student's choice of school and occupational field, and his attitudes in the areas of religion and politics. The questions were designed to elicit material concerning whether the student went through a period of crisis in arriving at his choice of occupation and belief systems and whether he was committed to them. For the purpose of the present study, only the material on occupation is relevant for testing the hypothesis.

For each S, the identity status material was classified by two judges. When disagreement as to the occupational identity status occurred, the interview was submitted to a third judge for decision.

Following the identity status interview, each student was questioned regarding his attitudes toward RPI. This material was also rated by two judges and was scored on a series of four 5-point scales: (a) satisfaction with their education, (b) satisfaction with the faculty, (c) satisfaction with the administration, and (d) satisfaction with the students. A high score indicated a favorable attitude, while a low score indicated a negative evaluation. When disagreements in ratings occurred, the two scores for that scale were averaged.

Results

Since there were no significant differences between freshmen and juniors in the frequency of the various occupational identity statuses or in

Table 1.—Mean Ratings for Satisfaction with College for Students with the Various Occupational Identity Statuses

	Satisfaction with			
Status	Education	Faculty	Administration	Students
Identity achievers (N=6)	4.17	3.42	3.08	2.67
Moratoriums (N=11)	3.05	2.82	3.05	2.91
Foreclosures (N=9)	4.44	3.61	3.67	3.28
Identity diffusions (N=6)	3.33	2.92	3.25	3.00

their evaluation of college, data from both classes were combined for all subsequent analyses.

The mean ratings for the student's satisfaction with college are found in Table 1. Analysis of variance reveals a significant difference between students with different occupational identity statuses regarding their evaluations of their education ($F = 5.24$, $df = 3/28$, $p < .01$). Consistent with the hypothesis, F's on occupation showed the most satisfaction while M's indicated the least satisfaction with their education. The results for satisfaction with the administration were in the same direction and approached significance ($F = 2.35$, $df = 3/28$, $p < .10$). No differences were observed for either satisfaction with the faculty or with the students at the school. Thus, the identity status of an individual regarding his occupational choice is an important factor in his evaluation of some, but not all, of his college experiences.

STUDY 2

In order to provide further evidence that the association of an identity crisis with college is an important factor in the student's evaluation of the school, previously obtained data from RPI seniors on the College Student Questionnaire–2 (CSQ–2) were reanalyzed. The CSQ–2 contains questions concerning the student's educational and vocational plans, his activities at college, his satisfaction with various aspects of the school, and his attitudes in a variety of areas. (For a more complete discussion of this questionnaire, see Peterson (7).) While the questionnaire did not provide enough information to assess the degree of crisis over occupation directly, it did provide sufficient material about the student's choice of major field to assess whether he went through a crisis in selecting his major field and, if he did, when it occurred.

There are two indicators of crisis during the college years which can be abstracted from the CSQ–2. The first concerns how and when the student selected his major field. Data from the questionnaire revealed that approximately 85 percent of the seniors in the sample were considering two or more alternatives when they chose their major field, i.e., they experienced a crisis. About half of these students reported that they made their decision in high school—HSD—while the remainder reported that the decision was made in college—CD. If it is the association between college experiences and the occurrence of crisis that leads to negative attitudes toward college, then it should be expected that CD would show less favorable evaluations of their school than would HSD. The second, and related, indicator of a crisis involving choice of major field was taken from the question on the CSQ–2 concerning whether or not the student changed his major field after entering college. Just over 40 percent of the students reported that they changed majors one or more times since entering college. Since change of major field is often indicative of crisis, it is hypothesized that students who changed fields will express less satisfaction with college than those who remained with the same major field throughout their college careers.

It must be emphasized that these indicators of crisis, the time of choosing between alternative majors and the changing of the major field, are not independent. A student who switched his major while in college is likely to have selected between alternative fields at that time. Therefore, these two measures may best be considered as alternative means of assessing the same underlying concept of crisis.

Method

Subjects. The Ss in this study were 282 RPI seniors from the Class of 1967.

Procedure. Each student completed the CSQ–2 during class time. The dependent measures were the four 10-item scales of the CSQ–2 relating to satisfaction with college: (a) satisfaction with the major field, (b) satisfaction with the faculty, (c) satisfaction with the administration, and (d) satisfaction with the students.

Results

The mean scores on each of the four satisfaction scales for HSD and CD are presented in Table 2. It can be seen that on all measures, CD show less satisfaction than HSD. This effect was significant for satisfaction with the major field ($t = 3.84$, $df = 235$, $p < .001$), satisfaction with the administration ($t = 2.31$, $df = 242$, $p < .05$), and satisfaction with the students ($t = 2.54$, $df = 239$, $p < .05$). The difference between groups regarding attitudes toward the faculty did not attain significance.

Table 3 contains the mean satisfaction scores for students who switched majors while in college and students who remained in the same field throughout their college careers. Here too, all the satisfaction scores are in the predicted direction. The t-test comparisons reveal that the differences are significant for satisfaction with the major

Table 2.—Mean Scores for Satisfaction with College for Students Deciding Between Alternative Major Fields While in High School and While in College

	Satisfaction with			
	Major	Faculty	Administration	Students
College Deciders	26.46	25.37	25.57	25.08
High School Deciders	28.62	26.20	27.07	26.40

Table 3.—Mean Scores for Satisfaction with College for Students Who Changed Their Major Field While at College and Those That Did Not Change

	Satisfaction with			
	Major	Faculty	Administration	Students
Changed	26.23	25.14	25.51	25.58
Did not change	28.61	26.37	27.22	25.90

(t = 4.63, df = 270, p < .001), satisfaction with the faculty (t = 2.23, df = 280, p < .05), and satisfaction with the administration (t = 2.88, df = 279, p < .01). The difference for satisfaction with the students was not significant.

Thus, for both indicators of crisis, those students who experienced a crisis while at college consistently exhibited more negative attitudes than did students who did not undergo a crisis.

Discussion

The findings from both studies demonstrate that the variable of ego identity status plays an important role in the student's reaction to college. Students who go through a moratorium concerning their choice of occupation or major field while at college have less favorable attitudes toward their school than students who did not go through an identity crisis during their college years. The results suggest that it is the association of the stress of crisis with college experiences, and not the fact of having had an identity crisis, that is the crucial factor in the relationship between ego identity and satisfaction with college. Support for this conclusion can be drawn from the data on students who had an identity crisis and resolved it before entering college. Since, for the HSD, the stress of crisis would not be associated with college, it was anticipated that they would have relatively favorable attitudes, and the findings bear this out.

It must be recognized that the experience of an identity crisis while at college is not thought to be indicative of psychological maladjustment. To the contrary, within the theoretical framework of Erikson and Marcia, the seeking of an identity from among available alternatives is an important precondition for further psychological growth. Therefore, this research does not imply that colleges should strive to reduce the probability of crisis. Rather, efforts should be directed toward providing an atmosphere which facilitates the student's successful resolution of his identity crisis.

FOOTNOTES

1. Study 1 was supported, in part, by funds contributed by the Office of the Vice-president for Student Affairs at Rensselaer Polytechnic Institute.
2. The authors wish to express their appreciation to Scott Bello and Henry Kurland who served as judges in Study 1, and to Robert Blase who aided in the data analysis for Study 2. A special note of gratitude is extended to Dr. Wynn Brown under whose auspices the data for Study 2 were collected.

REFERENCES

1. Beardslee, D. C.; O'Dowd, D. D., "Students and the Occupational World," in Sanford, N. (ed.), *The American College*, John Wiley and Sons, New York, 1962, pp. 597–626.
2. Douvan, Elizabeth; Withey, S. B., *A Study of Adolescent Boys*, Survey Research Center, The University of Michigan, Ann Arbor, 1955.
3. Erikson, E. H., "Identity and the Life Cycle," *Psychological Issues*, 1: Monograph 1, 1959.
4. Erikson, E. H., *Identity: Youth and Crisis*, Norton and Co., New York, 1968.
5. Marcia, J. E., "Development and Validation of Ego-identity Status," *Journal of Personality and Social Psychology*, 3:551–558, 1966.
6. Marcia, J. E., "Ego Identity Status: Relationship to Change in Self-esteem, 'General Maladjustment,' and Authoritarianism," *Journal of Personality*, 35:118–133, 1967.
7. Peterson, R. E., *Technical Manual: College Student Questionnaires*, Educational Testing Service, Princeton, New Jersey, 1965.
8. Summerskill, J., "Dropouts from College," in Sanford, N. (ed.), *The American College*, John Wiley and Sons, New York, 1962, pp. 627–657.
9. Trent, J. W.; Medsker, L. L., *Beyond High School: A Study of 10,000 High School Graduates*, Center for Research and Development in Higher Education, The University of California, Berkeley, 1967.

Forthcoming Articles

Gorton, Richard A.; University of Wisconsin–Milwaukee
"FACTORS WHICH ARE ASSOCIATED WITH THE PRINCIPAL'S BEHAVIOR IN ENCOURAGING TEACHER PARTICIPATION IN SCHOOL DECISION-MAKING"

Hakstian, A. Ralph; University of Massachusetts, Amherst
"THE EFFECTS OF TYPE OF EXAMINATION ANTICIPATED ON TEST PREPARATION AND PERFORMANCE"

Khan, Abdul Wahab; West Pakistan University of Engineering and Technology, Lahore
"TIMETABLE IN A TECHNICAL UNIVERSITY"

Chang, Sunnyuh Shin; Raths, James; University of Maryland, College Park
"THE SCHOOLS' CONTRIBUTION TO THE CUMULATING DEFICIT"

Smith, Leland; University of Kentucky, Lexington
"A 5–YEAR FOLLOW-UP STUDY OF HIGH ABILITY ACHIEVING AND NONACHIEVING COLLEGE FRESHMEN 1960–1966"

Taylor, Hugh; University of Victoria, British Columbia
"STUDENT REACTION TO THE GRADE CONTRACT"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Relations Between Grade Point Averages and Collegiate Course Grade Distributions[1]

WILLIAM A. MEHRENS
Michigan State University

BRUCE G. ROGERS
University of Maryland

ABSTRACT

Many educational institutions establish a uniform grade distribution policy and encourage its general observance. Most schools also establish minimum grade point average (GPA) restrictions for participating in various activities. Unfortunately, these policies are often made independently of each other. In this paper, a mathematical model depicting a relationship between the distributions of course grades and the GPA distribution is proposed. To empirically test the proposed model, course grade data were obtained from 837 undergraduates. The predicted and actual GPA distributions were in reasonable agreement. The model may prove useful in those settings where it is desired to predict the effect of proposed classroom grading policy on the GPA distribution prior to the actual implementation of that policy.

MEASUREMENT specialists, in general, advocate the use of relative marking systems in preference to absolute systems. But since relative systems typically permit the instructor to control the percentages of A's, B's, etc., to be awarded, it is possible for these proportions to vary widely between classes, even for groups of similar abilities. Accordingly, many institutions seek to establish a uniform grade distribution policy and encourage or require adherence to it by the faculty.

The desire for grade distribution uniformity is understandable in light of the many important decisions based on GPA. Most secondary schools and colleges have minimum GPA restrictions for participation in various activities, e.g., athletics and honors courses. At the undergraduate and graduate levels of college, minimum GPA restrictions often control the admission and expulsion of students. It, therefore, seems reasonable that when an institution desires to modify its classroom grading policies, it should begin by assessing the probable effects that will result in the GPA distribution.

No published studies presenting either theoretical formulations or empirical data concerning the relationship between grade distributions and GPA distributions were found. The present investigation was initiated to explore the interrelationships between the distributions by proposing and empirically testing a simple mathematical model.

The Model

Assume $x_1, x_2, \ldots x_K$ are random variables, each multi-normally distributed, and let

$$y = \frac{1}{K}\sum_{i=1}^{K} x_i \tag{1}$$

From a theorem in multivariate analysis (1:25) it follows that y will be normally distributed with

$$E\ (y) = \frac{1}{K}\sum_{i=1}^{K} E\ (x_i) \tag{2}$$

and

$$\sigma_y^{\ 2} = \left(\frac{1}{K}\right)^2 \left[\sum_{i=1}^{K} \sigma_{x_i}^{\ 2} + 2 \sum_{i<j}\sum r_{x_i x_j} \sigma_{x_i} \sigma_{x_j} \right] \tag{3}$$

Impose the added restrictions of equal variances of the K sets, and let this common value be σ_x^2. Set ρ equal to the average intercorrelation of the K sets

$$\bar{\rho} = \frac{2 \sum_{i<j}\sum r_{x_i x_j}}{K\ (K - 1)} \tag{4}$$

The expression for σ_y^2 now simplifies to

$$\begin{aligned} \sigma_y^{\ 2} &= \frac{1}{K^2}\left[K \sigma_x^{\ 2} + K\ (K - 1) \bar{\rho}\ \sigma_x^{\ 2} \right] \\ &= \frac{\sigma_x^{\ 2}}{K} \left[1 + (K - 1) \bar{\rho} \right] \end{aligned} \tag{5}$$

Consequently knowledge of ρ, σ_x^2, and E (x_1), along with the assumption that the variables are normally distributed, is sufficient to describe the distribution of y.

Comparison of the Model with Actual Data. With some exceptions, all undergraduate students at Michigan State University are required to take four course sequences of three terms each, including American Thought and Language (ATL), Natural Science (NS), Social Science (SS), and Humanities (HUM). The first two are generally taken during the freshman year and the last two in the sophomore year.

Grades were obtained on all those students who enrolled in the fall of 1958 and completed the twelve terms of four basic courses during the next 4 years. (The resulting sample of 837 students probably yielded somewhat restricted distributions, since the poorer students would be more likely to withdraw from school while the better students would be more likely to use waiver examinations.) The mean and variance for each variable are presented in Table 1. The intercorrelations of the course grades are shown in Table 2. They ranged from .197 to .638, with an average value of .39.

Table 1.—Means and Variances of Final Grades (A = 4) in a Sample of 837 Undergraduates (See Text for Course Titles)

Variables	Mean	Variance
1. ATL 111	2.48	.486
2. ATL 112	2.55	.591
3. ATL 113	2.49	.580
4. NS 181	2.43	.558
5. NS 182	2.42	.578
6. NS 183	2.36	.792
7. SS 231	2.48	.642
8. SS 232	2.28	.750
9. SS 233	2.28	.759
10. HUM 241	2.43	.606
11. HUM 242	2.35	.693
12. HUM 243	2.34	.663
13. GPA	2.41	.284

Since the data do not precisely meet the assumptions of equal variances among courses, the value of σ_x^2 will be estimated as the mean variance of the twelve grade distributions.

$$\hat{\sigma}_x^2 = \frac{1}{12}\sum_{i=1}^{12}\left(\sigma_{x_i}^{\;2}\right) = .64 \qquad (6)$$

Using this value and $\rho = .39$ (from formula (4)), the predicted GPA variance, according to formula (5), is

$$\hat{\sigma}_{GPA}^2 = .282$$

The difference between this value and the actual GPA variance of .284 (from Table 1) appears to be negligible for practical purposes.

In Table 3, the frequency distribution of the GPA values is divided into ½ grade point units to compare the predicted and actual proportions in each category (2:485–7) using the Kolmogorov–Smirnov test (3:47–52). The maximum deviation between the observed and predicted cumulative frequency, called D, is .028, which is not significant even at the .20 level ($D_{.20} = .037$). This accuracy, however, may be somewhat inflated by using the same data to generate both the predicted and observed values. Unfortunately, it was not possible to obtain another sample for cross validation.

Discussion

At the beginning of this paper it was suggested that the relationship between classroom grade distributions and GPA distributions might be predictable from a simplified model. An empirical test of the model showed that the assumption of common variance and common intercorrelations of classroom grade distributions yielded a predicted GPA distribution remarkably close to the actual GPA distribution. For example, the theoretical model predicted that approximately 22 percent of the students would fail to have a GPA of at least 2.0, compared with 21 percent in the actual data.

It may be questioned why one would go to all the trouble of predicting the GPA distribution from course grade distributions when the GPA distribution itself can be computed directly therefrom. The answer is that an educational institution may desire to assess the probable effect of a proposed institutional grading policy, *prior* to the actual implementation of the policy. The results presented herein indicate that the errors in

Table 2.—Intercorrelations of the 13 Variables (Final Grades) on 837 Undergraduates*

Course No.	ATL 111	ATL 112	ATL 113	NS 181	NS 182	NS 183	SS 231	SS 232	SS 233	HUM 241	HUM 242	HUM 243
ATL 112	581											
ATL 113	566	624										
NS 181	362	409	367									
NS 182	363	434	413	526								
NS 183	197	273	226	492	500							
SS 231	378	397	482	306	356	219						
SS 232	233	264	290	287	345	352	521					
SS 233	242	331	349	270	330	263	589	581				
HUM 241	376	405	471	362	422	319	426	407	438			
HUM 242	365	410	454	329	391	262	413	397	432	638		
HUM 243	369	407	476	306	398	261	413	362	429	599	611	
GPA	612	681	704	620	681	555	689	643	668	732	715	704

*The decimal points have been omitted.

Table 3.—Actual and Predicted Proportions of the GPA Distribution Intervals of ½ Grade Point

GPA Interval[a]	Actual Proportion	Predicted Proportion	Cumulative Actual Proportion Minus Predicted Proportion
(≥3.5)	.0347	.0202	
(3, 3.5)	.1254	.1433	-.0146
(2.5, 3)	.2652	.2690	.0033
(2, 2.5)	.3620	.3469	.0071
(1.5, 2)	.1971	.1770	-.0080
(1, 1.5)	.0143	.0395	-.0281
(< 1)	.0012	.0041	-.0029

[a] The interval begins with the bottom value and runs to, but not including, the top value.

prediction from the proposed model are not likely to be of practical significance. In this context, it is hoped that the model may prove of use in the appropriate decision-making processes.

FOOTNOTE

1. This report is based on a paper presented at the 1968 Convention of the American Educational Research Association, Chicago, Illinois.

REFERENCES

1. Anderson, T. W., *An Introduction to Multivariate Statistical Analysis,* Wiley, New York, 1958.

2. Peters, C. C.; Van Voorhees, W. R., *Statistical Procedures and Their Mathematical Bases*, McGraw-Hill, New York, 1940.

3. Siegal, Sidney, *Nonparametric Statistics,* McGraw-Hill, New York, 1956.

JeR Research Briefs

A NOTE ON TWO CPI VARIABLES AS PREDICTORS OF NONACADEMIC ACCOMPLISHMENT IN COLLEGE

ACADEMIC accomplishment, i.e., getting good grades, may be thought of as "Structured Achievement," or effective behavior in the classroom environment (a situation requiring conformity). Nonacademic (extracurricular) accomplishment may be described as "Self-initiated Achievement," or effective self-expression which occurs outside required classroom work as a result of independent endeavor (2). In constructing the California Psychological Inventory (CPI), Gough (1) distinguished two types of achievement tendency in terms of personality variables:

(Continued on page 189)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Effects of Certain Audio and Visual Presentation Sequences on Pupil Information Acquisition

ARMAND J. GALFO
College of William and Mary, Williamsburg, Virginia

ABSTRACT

The purpose of the investigation was to determine if (1) pupils gain more information from slide-tape audio-visual materials when sight and sound are not presented simultaneously; and (2) pupil information acquisition is influenced by whether the audio or visual message is presented first. Randomly selected groups of pupils were exposed to two experimental sequences which separated sight and sound and two control treatments which combined the audio and visual presentation. An original hypothesis, based on previous studies of human perception, that sight–sound separation would produce better pupil information acquisition than the control conditions of simultaneous sight–sound, was abandoned after a pilot study. The reverse hypothesis, that simultaneous sight–sound would prove superior to sight–sound separation, was supported.

THE PURPOSE of the study was to determine if different methods of sequencing audio and visual portions of a slide-tape presentation would result in differences in pupil information acquisition. Specifically, answers to the following questions were sought:

(1) Do pupils learn more from the given audio–visual presentations if sight and sound are not presented simultaneously?

(2) When audio messages and visuals are separated, does pupil information acquisition vary as a result of the audio or visual message being placed first in the presentation sequence?

Two experimental and two control sequences, using slide-tape presentations, were studied:

(1) An experimental mode in which each bit of audio information was followed by the corresponding visual frame (A→V).

(2) An experimental mode in which each visual frame was followed by the corresponding audio (V→A).

(3) A control mode in which audio–visual frames occurred simultaneously as in a normal audio–visual lesson but with pauses between the frames to give a total time for the lesson equal to the first two types of sequences (A–V pause).

(4) The normal audio–visual control sequence (A–V) in which audio–visual information frames followed one another in succession

The purpose served, and specific objectives sought by the experiment were associated with the whole question of perception which has been studied by psychologists for many years. The function of set in perception, for example, is discussed extensively in the literature of psychology. This is also true of the act of attending. In summarizing such studies, Solley and Murphy (5) point out that: "There can be little doubt that what we perceive is (in part) determined by . . . the potentiation of one set of factors by another." If this is accepted, then it is logical to hypothesize that the sequence in which individuals receive audio and visual stimulation may have some bearing upon how much they perceive; and therefore, how much they learn. In other words, sequence of communication modes may be considered to be "factors" which "potentiate."

The notion that various methods of sequencing audio–visual messages may bring about differences in learning seems to fit the Broadbent (1) model of the human perceptual system. The model, used by Broadbent and others to guide experiments in perception, represents the human receiver of information as one which operates with but one communication channel open at a given time. Travers and others (6) in reviewing research and theory related to information transmission through audio–visual methods, concluded ". . . the Broadbent model suggests that auditory word presentation should not occur simultaneously with the visual object presentation. . . ." While the model does not indicate whether succession should be audio–visual or visual–auditory, other information from experimental psychology suggests that order may be important.

From the theoretical considerations which have been discussed, two hypotheses were originally developed to guide the study. The Broadbent model of perception, and studies by Lumsdaine and others on audio–visual cuing, suggested that:

1. The experimental modes (A→V and V→A), which separate sight and sound, would result in greater pupil information acquisition than the control modes (A–V pause and A–V) which present the visual and audio message simultaneously.
2. One of the experimental modes would prove superior to the other in promoting pupil information acquisition.

A pilot study, however, indicated that the reverse of hypothesis one might be true; i. e., control conditions with unseparated sight–sound would produce superior information acquisition. The pilot study also produced conflicting results regarding the second hypothesis. In one experiment V→A mode was superior to A→V, but in another experiment neither mode was superior.

It was hypothesized, therefore, that if pupils were presented lessons in the sequence modes described as A→V, V→A, A–V pause and A–V.

1. The control modes (A–V pause and A–V, which present sight and sound together would result in greater pupil information acquisition than the control modes (A→V and V→A) which separate sight and sound.
2. Groups exposed to the two control modes would not differ significantly in their acquisition of information.

It was also decided to retest the original hypothesis two but with a prediction that V→A sequence would produce better results than A→V.

Procedure

For the final experiment, filmstrips were appraised in terms of three criteria: (1) suitability for the junior high school grades; (2) subject matter not familiar to the groups who were to take part in the study; and (3) information presented in approximately equal proportion as strictly audio, strictly visual, and both audio and visual. The lesson selected was part of a series on the geography of Africa—*The Eastern Highlands*— produced by the Society for Visual Education. It consisted of sixty-three information frames.

Students who were to take part in the experiment were selected from the eighth and ninth grades in two school systems which had not been involved in the pilot study. The mental ability of all eighth and ninth graders in the two junior high schools selected were recorded and separated into three categories: high ability (117 or higher IQ), average ability (90–110 IQ), and low ability (85 or lower). Since the criterion test was to be administered in written form, pupils with a reading level below grade four were excluded. Forty-four high ability students, forty-four of average ability, and forty-four of low ability from each grade were selected at random to take part in the study and assigned at random to the treatment groups as shown in Table 1.

Factors such as the time of day when lessons were presented and the physical setting were carefully controlled. Also, the audio for all sequence modes was originally made into a master tape by a professional announcer; the tapes for the various lessons were thereby the same except in terms of the sequence mode of presentation.

Table 1.—Experimental Design—Final Experiment

Ability Level	Treatment Mode							
	A→V		V→A		A-V (pause)		A-V	
	grade		grade		grade		grade	
	9th	8th	9th	8th	9th	8th	9th	8th
High	11	11	11	11	11	11	11	11
Average	11	11	11	11	11	11	11	11
Low	11	11	11	11	11	11	11	11

N = 264

The evaluation instrument consisted of a 120-item multiple choice test. Forty items came from the audio source of information, forty from visual source of information and forty were selected in terms of information that was presented by both sight and sound. (For obvious reasons, however, the various types of items were mixed together at random throughout the test.) The test was prepared in this 3-fold form as an attempt to control and analyze pupil information acquisition in terms of the specific percept source or sources.

Findings

The pupil test scores which resulted from the final experiment were first subjected to an analysis of variance in terms of six sets of scores: (1) total scores (audio information + visual

Table 2.—Analysis of Variance—Total Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	941	3	314	2,77*
Ability Level	L	26202	2	13101	115,94**
Grade	G	1201	1	1201	10,63**
L x G Interaction		791	2	396	3,50*
L x T Interaction		757	6	126	1,12
G x T Interaction		645	3	215	1,90
L x G x T Interaction		622	6	104	92
Error		27135	240	113	

* Significant at the 5% level
** Significant at the 1% level

Table 3.—Analysis of Variance—Visual Information Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	133	3	44	2.00
Ability Level	L	4138	2	2069	94.05**
Grade	G	130	1	130	5.90*
L x G Interaction		84	2	42	1.91
L x T Interaction		142	6	24	1.09
G x T Interaction		104	3	35	1.59
L x G x T Interaction		103	6	17	.77
Error		5359	240	22	

Table 4.—Analysis of Variance—Audio Information Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	52	3	17	.85
Ability Level	L	3608	2	1804	90.20**
Grade	G	131	1	131	6.55*
L x G Interaction		90	2	45	2.25
L x T Interaction		104	6	17	.85
G x T Interaction		59	3	20	1.00
L x G x T Interaction		94	6	16	.80
Error		4769	240	20	

* Significant at the 5% level
** Significant at the 1% level

Table 5.—Analysis of Variance—Audio–Visual Information Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	188	3	63	3.70*
Ability Level	L	1224	2	612	36.00**
Grade	G	89	1	89	5.24*
L x G Interaction		80	2	40	2.35
L x T Interaction		88	6	15	.88
G x T Interaction		85	3	28	1.65
L x G x T Interaction		122	6	20	1.17
Error		4193	240	17	

* Significant at the 5% level
** Significant at the 1% level

Table 6.—Analysis of Variance—A-V + Visual Information Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	585	3	195	3.42*
Ability Level	L	10707	2	5398	94.70**
Grade	G	472	1	472	8.28**
L x G Interaction		322	2	161	2.82
L x T Interaction		343	6	57	1.00
G x T Interaction		339	3	113	1.98
L x G x T Interaction		367	6	61	1.07
Error		13736	240	57	

* Significant at the 5% level
** Significant at the 1% level

Table 7.—Analysis of Variance—A-V + Audio Information Scores

Source of Variation		Sums of Squares	df	Mean Square	F-ratio
Treatment	T	357	3	119	2.47
Ability Level	L	9798	2	4899	102.06**
Grade	G	472	1	472	9.83**
L x G Interaction		283	2	141	2.94
L x T Interaction		255	6	43	.90
G x T Interaction		173	3	58	1.21
L x G x T Interaction		335	6	56	1.17
Error		11422	240	48	

* Significant at the 5% level
** Significant at the 1% level

Table 8.—Orthogonal Comparisons of Treatment Means—Total Scores

Source of Variation	Treatment Means				
	A—►V 56.70	V—►A 55.59	A-V (Pause) 58.98	A-V 60.41	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A—►V)-(V—►A)	40	4%	1	40	.35
C_2 : (AVP)-(AV)	67	7%	1	67	.59
C_3 : (Cont. - Exp.)	834	89%	1	834	7.38**
Treatments	941	100%	3	314	2.77*
Error term	27135		240	113	

* Significant at 5% level
** Significant at 1% level

information + audio–visual information); (2) visual information scores; (3) audio information scores; (4) audio–visual information scores; (5) visual + audio–visual information scores; and (6) audio + audio–visual information scores. Tables 2 through 7 summarize the results of the analyses of variance. (The error term used throughout was the within-groups mean square since the experimental design was considered to be a fixed model).

Although main effects F-ratios for solely audio and solely visual information scores were not significant, Tables 2, 5, and 6 show that there were significant differences in the treatment means for the other scores analyzed. A significant interaction appeared in the experiment in the levels × grade factor for total scores.

In order to ferret out the sources of main effect differences, and to permit more precise and appropriate tests of predictions, orthogonal com-

Table 9.—Orthogonal Comparisons of Treatment Means—Visual Information Scores

Source of Variation	Treatment Means				
	A→V 20.35	V→A 20.33	A-V (Pause) 21.79	A-V 21.73	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A→V)-(V→A)	.02	negligible	1	.02	.0009
C_2 : (AVP)-(AV)	.12	negligible	1	.12	.005
C_3 : (Cont. - Exp.)	132	almost 100%	1	132	6.00*
Treatments	133	100%	3	44	2.00
Error term	5359		240	22	

* Significant at 5% level
** Significant at 1% level

Table 10.—Orthogonal Comparisons of Treatment Means—Audio Information Scores

Source of Variation	Treatment Means				
	A→V 17.79	V→A 17.00	A-V (Pause) 18.23	A-V 17.80	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A→V)-(V→A)	20	38%	1	20	1.00
C_2 : (AVP)-(AV)	6	12%	1	6	.30
C_3 : (Cont. - Exp.)	26	50%	1	26	1.30
Treatments	52	100%	3	17	.85
Error term	4769		240	20	

* Significant at 5% level
** Significant at 1% level

Table 11.—Orthogonal Comparisons of Treatment Means—Audio-Visual Information Scores

Source of Variation	Treatment Means				
	A→V 18.50	V→A 18.26	A-V (Pause) 19.50	A-V 20.36	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A→V)-(V→A)	5	2%	1	5	.29
C_2 : (AVP)-(AV)	25	14%	1	25	1.47
C_3 : (Cont. - Exp.)	156	84%	1	156	9.18**
Treatments	186	100%	3	63	3.70*
Error term	4193		240	17	

* Significant at 5% level
** Significant at 1% level

Table 12.—Orthogonal Comparisons of Treatment Means—Visual + Audio-Visual Information Scores

Source of Variation	Treatment Means				
	A→V 38.74	V→A 38.59	A-V (Pause) 40.98	A-V 42.09	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A→V)-(V→A)	1	negligible	1	1	.02
C_2 : (AVP)-(AV)	40	7%	1	40	.70
C_3 : (Cont. - Exp.)	544	93%	1	544	9.54**
Treatments	585	100%	3	195	3.42*
Error term	13736		240	57	

* Significant at the 5% level
** Significant at the 1% level

Table 13.—Orthogonal Comparisons of Treatment Means—Audio + Audio-Visual Information Scores

Source of Variation	Treatment Means				
	A→V 36.33	V→A 35.26	A-V (Pause) 37.48	A-V 38.33	
	SS	% of treatment SS due to source	df	MS	F-ratio
C_1 : (A→V)-(V→A)	38	10%	1	38	.79
C_2 : (AVP)-(AV)	24	7%	1	24	.50
C_3 : (Cont. - Exp.)	295	83%	1	295	6.14*
Treatments	357	100%	3	119	2.47
Error term	11422		240	48	

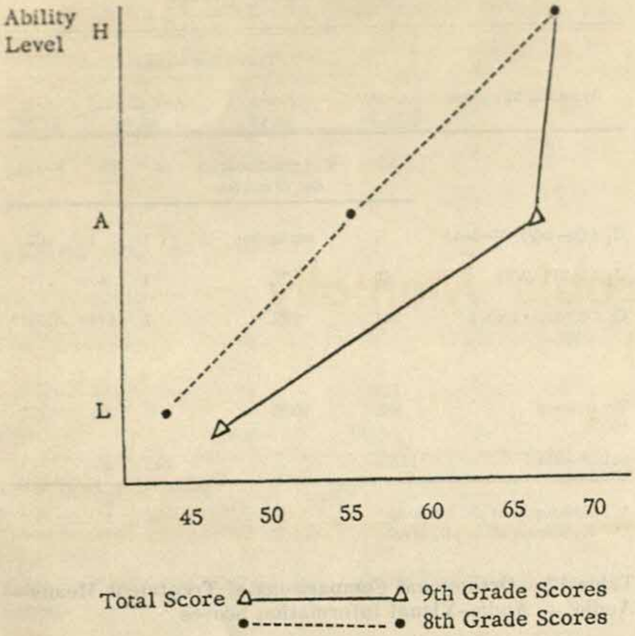
* Significant at the 5% level
** Significant at the 1% level

parisons of main effects means were made. These comparisons are summarized in Tables 8 through 13.

The graph of levels × grade interaction effect that showed up in the total scores indicates quite clearly that the high ability eighth and ninth graders scored equally well. (See Figure 1.) The low ability eighth and ninth graders' scores were also fairly equal. The average ability groups' scores from the two grades, however were quite different.

The guidance provided by the pilot study and preliminary experiments in suggesting the alternative hypotheses to be tested proved very useful in that planned comparisons could be made which supported two of the hypotheses. The orthogonal comparisons, showed that even where overall F-ratios across all experimental and control treatments was not sufficiently high to reject null at the 5 percent level of confidence, control groups

Figure 1—Levels × grade interaction effect—Total Scores



(simultaneous AV scores) were significantly higher than experimental groups (sight–sound separation scores) in all save the solely audio information scores.

Every orthogonal comparison between control groups, furthermore, supported the second alternative hypothesis that the two control modes (AV pause and A–V) would not perform differently; this in spite of the fact that several of the overall F-ratios *were* significant. The result also tends to give additional support to alternative hypothesis one since the control modes have simultaneous audio–visual messages in common but differ in the length of the lesson presentation. (If length of time in the lesson had affected the controls differently, using their scores in combination to check hypothesis one would have been a questionable procedure.) The hypothesis that one of the experimental modes would prove superior to the other was not supported by the data.

Limitation of the study in terms of population sampled and the particular materials used in the study leaves several unanswered questions. For instance, would other age groups perform similarly with respect to the audio–visual sequence modes used? Would the use of other audio–visual materials produce the same results? Did the use of different materials in the pilot study and final experiment contribute to conflicting results regarding the predicted variability of experimental sequence modes? These questions would suggest a need for new experiments which consider hypothesis one with different populations and materials.

Conclusion

Conclusions cannot be drawn regarding the reasons why the original hypothesis which predicted superiority of the sight–sound separation had to be rejected in favor of its reverse. Several possibilities suggest further research, however. It is entirely conceivable, for example, that the Broadbent theory is correct but that human attention switches so rapidly from one communication channel to another that any attempt to force an individual into an externally regulated pattern reduces, rather than enhances, information acquisition. It may also be that the Broadbent theory is incorrect and that people do perceive information from several communication channels at once. The results of the study neither support nor refute these two notions.

In either event, however, the research which was conducted suggests very strongly that in terms of practical use with audio–visual materials, sight–sound separation is not an effective method. If the second possibility suggested is true—that the Broadbent theory is incorrect—then obviously sight–sound separation, as a way of developing more informative audio–visual materials for children like those who took part in the experiment, will not work. But even if the first possibility is true—that Broadbent is right, but that human attention switches very rapidly from one communication channel to another—it would appear that any attempt to take advantage of the phenomenon through artificial sight–sound separation of audio visual presentations will be met with frustration.

Finally, although one of the pilot experiments showed a significant difference in information acquisition by an experimental (sight–sound separation) group over another, another pilot experiment and also the final experiment did not show the same result. This would warrant a test of a hypothesis (using different A–V lessons) that neither A→V nor V→A sequence mode will prove superior.

REFERENCES

1. Broadbent, D. E., *Perceptions and Communication*, Pergamon Press, New York, 1958.
2. Lumsdaine, A. A., "Instruments and Media of Instruction" in Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand McNally & Co., Chicago, 1963, Chapter 12, pp. 670–671.
3. May, M. A.; Lumsdaine, A. A., *Learning From Films*, Yale University Press, New Haven, Connecticut, 1958.
4. Nue, D. M., *The Effect of Attention–Gaining Devices on Film–Mediated Learning*, Office of Naval Research, Port Washington, New York, 1950.
5. Solley, Charles M.; Murphy, Gardner, *Development of the Perceptual World*, Basic Books, New York, 1960.
6. Travers, Robert M. W.; McCormick, Mary C.; Van Mondfrans, Adrian P.; Williams, Frank E., *Research and Theory Related to Audio–Visual Information Transmission*, U. S. Department of Health, Education and Welfare, Washington D. C., 1964.
7. Travers, Robert M. W., "A Study of the Relationship of Psychological Research to Educational Practice, in Glaser R., (ed.), *Training Research and Education*, Pittsburgh.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Teaching Arithmetic to Educable Mentally Retarded Children (Review)

THOMAS NOFFSINGER
Lake County Regional Program for Handicapped Children, Mentor, Ohio

VIRGINIA DOBBS
Pupil Personnel Director
Metropolitan Nashville Public Schools, Tennessee

ABSTRACT

Research related to the arithmetic ability of mentally retarded children has evolved into a great diversity of special learning areas. This review categorizes eight special interest areas and highlights the unique contributions of each. Conclusions point to the continued need for research in each area as well as several major findings.

THE IDENTIFICATION of characteristics of the mentally retarded child's arithmetic ability has long been a topic of research. However, interested scholars have had to scan journals from many fields to keep abreast of the progress being made in this area of study. This review has been compiled to provide, in one place, eight major areas of interest in the arithmetic ability of educable mentally retarded (EMR) children and samples of research in those interest areas. The interest areas have been arbitrarily arranged in order to organize a highly diversified body of literature although the interest areas are not mutually exclusive.

The purpose of this review is not to identify specific studies in arithmetic research for special recognition, rather it is intended to be comprehensive in nature, dealing with many different topics of research in an effort to stimulate further interest in an area which offers good research potential as well as success experiences for EMR children.

General Characteristics

There has been much research which attempted to describe the general characteristics of the mental retardate's arithmetic ability. It has been known for decades that retardates' math ability is less than normal (4, 59). Costello's dissertation (18) tried to specify the particular types of deficiencies and their possible remedies. Costello concluded that the greatest arithmetic gains were made under social and concrete circumstances and stated that instruction related to social need was most effective. Young (92) continued Costello's plea that math be directed toward meeting social skills and job placement. Carrison and Werner (15) found that retardates were capable of arithmetic gains. Cruickshank's classic works (21, 22, 23) which compared the performance of the retarded to the non-retarded resulted in the following conclusions: Retardates are (1) significantly less able to solve the needed facts for a problem from superfluous material, (2) significantly less able to solve arithmetic problems when given the necessary facts in verbal form, (3) more nearly like the non-retarded child in straight arithmetic computation, (4) more primitive than normal children in habits, such as counting on their fingers, (5) more careless and prone to make more errors than normal children, (6) more prone than normal children to have difficulty naming the processes involved in problem solving.

In the Illinois Study (34), Goldstein, Moss, and Jordan tried to overcome the often observed discrepancy between the EMR's arithmetic reasoning ability and his computation ability by teaching the experimental group academic subjects as a means of developing abilities in problem solving. They found that the low intelligence quotient (IQ) experimental children were significantly superior to the control group in arithmetic computation *and* problem solving. There was no significant difference in the high IQ group. These results support the idea that mental retardates can learn reasoning skills if the curriculum is directed toward that goal.

Klausmeier, Feldhusen, and Check (48), in an important study of the arithmetic abilities of twenty boys and girls in each of three IQ levels (low, 50–80; average, 90–100; high, 120+) concluded that: (1) Speed of acquiring arithmetic learning is not inversely related to retention among groups of differing intellectual abilities. (2) The ratio of relearning time is the same among children of low, average, and high ability. However, the rate of acquiring overall arithmetic learning is much lower for low IQ than the other two groups. (3) Retention of arithmetic learning is the same among children when original learning is graded to the learner's achievement level. (4) Uneven physical development is not significantly related to low efficiency in arithmetic achievement. (5) Within child variability in the low and high groups was lower than the average group.

Jarvis (44) found that longer (60 minutes) class periods were substantially related to better arithmetic achievement when contrasted to shorter periods whether the pupils were dull, average, or bright. Brown (11), in a study with institutional retardates, indicated that even his best group could not solve practical arithmetic although they had attained mechanical arithmetic ability to solve simple multiplication and division. Thus, from a descriptive point of view, the general characteristics seem well defined.

Concept Formation

Concept formation research has been extensively reviewed (29, 82) and the reader is referred to these two volumes for a detailed analysis. It is assumed that much of this research is indirectly applicable to teaching arithmetic, but there have been several studies in this area dealing directly with arithmetic concepts and the mentally retarded. Woodward (91) found that similar types of thinking as found in normal children have been observed in subnormal groups. She states that Piaget's "ascending hierarchy of cognitive processes" is a model applicable to arithmetic and the mentally retarded. This model states that concept formation is matured in stages which can be observed separately and are cumulative in nature. The absence of maturation in one stage would inhibit the presence of the following stages. These stages, furthermore, are a product of physical and mental growth and cannot be prematurely advanced. Jean Piaget (67) states, "It is a great mistake to suppose that a child acquires the notion of number and other mathematical concepts just from teaching. On the contrary, to a remarkable degree he develops them himself, independently and spontaneously. When adults try to impose mathematical concepts on a child prematurely, his learning is merely verbal; true understanding of them comes only with his mental growth."

Verification of the theory of concept formation in developmental stages has come from all over the world. Chase (17) in America, Estes (30) and Woodward (91) in the British Isles, Wang, and others (87) in Japan, Linhart (52) in Russia, and, of course, Piaget (67) in Switzerland are just a few of many investigators in this area of arithmetic. These studies corroborate that arithmetic concepts can be classified rightly to correspond to certain mental ages. The problem for research at the present time is to develop arithmetic teaching methods to correspond to each particular concept formation stage.

The importance of concept formation to EMRs has been noted by Hood (39). He compared forty mental subnormals and 126 normal children equated on mental age (MA) (7.8 – 9.6) on arithmetic development. He concluded that the retardates' failure in arithmetic was due to the absence of pre-number concepts, and that in spite of their MA, they were were not mentally ready for arithmetic. They were still at Piaget's intuitive stage of thinking and had not reached the stage of concrete observations. Hood went on to note that many children who fail Piaget tests may, nevertheless, be drilled so expertly in mechanical arithmetic that they appear to have skills and that their real lack comes in problem solving. This study is in harmony with EMRs better ability in computation than in reasoning (46).

Organicity Implications

Finger Localization. In the late 1930's, as a result of the body image studies of Schilder (75), a great emphasis in research was put on what we now might call motor and perceptual impairments. Strauss and Werner (83) made the contention that an organic deficiency could cause differential achievement abilities. They also noted that there was a positive relationship between arithmetic ability and the ability of children whose fingers were hidden to identify which of their fingers was being touched. Werner and Carrison (89) developed a finger schema test which correlated positively with arithmetic achievement. This idea supported the further studies of Strauss and the educational techniques of Lehtinen (50) with brain-injured children. Benton, and others (6) found no support for the Strauss and Werner ideas since both their normal and defective children demonstrated no association between arithmetic ability and finger localization or right-left discrimination. At the present time it seems doubtful that the finger localization theory has any direct relevancy for the educator, but there are still investigators finding positive correlations (54).

Endogenous vs. Exogenous. The relationship of etiology and arithmetic achievement also came out of the 30's and 40's with Lehtinen and Strauss (50) as the major exponents. Bensberg (5) found etiology to have a nonsignificant relationship with achievement. Two major studies of the 50's have dominated the research in this area. The first is the classic Capobianco (14) which found no significant differences between endogenous and exogenous subjects (Ss) on computation, arithmetic reasoning, and general areas of arithmetic achievement. He also concluded no significant differences qualitatively in terms of rigidity, reversals, and the concept of zero. The second study, done by Gallagher (33), found no significant differences on measures of quantitative concepts between brain injured and non-brain injured mentally retarded children. These studies seem to present strong evidence that there is little difference in arithmetic ability between endogenous and exogenous retardates. However, there is still some research on the topic which indicates that the question is not settled. Voerg (86) found that his mentally defective (organic) Ss did not perform as well as an MA matched group of mental retardates and a group of normals.

Mental Age and Arithmetic Computation

The research on the relation of arithmetic achievement to MA has been summarized by Kirk (46). He states that, "Retarded children appear to achieve in harmony with their mental age in arithmetic fundamentals but below their mental age expectancy in arithmetic reasoning problems requiring reading." He bases this conclusion primarily on Dunn's (27) conclusion but also cites: (1) Cruickshank (22, 23), who found no significant difference between normal and retarded in arithmetic gains. In these studies Cruickshank felt that arithmetic age corresponded to mental age; (2) Ring (74), who found mental retardates working at MA expectancy in arithmetic computation; and (3) Bensberg (5), who found institutional retardates with a mean arithmetic achievement within one month of their MA.

Quay (71), in a review of academic achievement of retardates, cites these findings on arithmetic and MA expectancies: (1) Merrill (59), found special class retardates below MA expectancy in all areas except arithmetic fundamentals. (2) Wilson (90) found his dull group superior on arithmetic achievement to his normal group matched on MA. (3) Jacobs (42) concluded that although mean total achievement scores were approximately equal to mean expected scores, his subjects were significantly higher in arithmetic than reading and language and significantly higher in arithmetic fundamentals than in arithmetic reasoning. Quay concluded that (71:681), "On the basis of the size of discrepancy between actual achievement and expected achievement calculated from MA, the order of skills is as follows: reading, arithmetic reasoning, writing, spelling, and arithmetic computation."

The research on the expected MA achievement is clouded with many variables such as differences in IQ levels of the groups under study, differences in methods of comparisons, and even differences in measurement of MA. However, there have been several studies which show arithmetic gains and could support the idea that MA and arithmetic computation are good achievement predictors. Capobianco (14) concluded that mentally handicapped children are capable of achievement on arithmetic up to their MA. He further concludes that MA appears to be more important than etiology in respect to qualitative functions. Bennett (4) found mentally retarded in special classes and regular classes achieving most in arithmetic computation. Blatt (8) found significant gains in arithmetic achievement. Thurstone (85) also found significant gains in arithmetic achievement and computations. She further stated that the mentally handicapped came nearest their predicted score in arithmetic computation which indicated they can profit more from drill and rote memory than from materials which involve more thinking and reasoning. Mullen and Itkin (63) state that their educable mentally handicapped sample, as a whole, gained less on arithmetic than might be expected of their IQ. They also noted that the only significant difference between the special class and the regular class EMR's was a superior arithmetic gain by the regular class in the first year. Finley (31) reported no difference in arithmetic achievement between retarded and normal of the same MA. These consistent gains in arithmetic for EMR's indicate an area of study where success experiences may be attained.

Learning Research

There have been many studies on MA and learning processes of the retarded. The reader is directed to Denny (24) and several chapters in Ellis (29) for excellent reviews of learning research. A few results of similar studies having direct implications for arithmetic are reported here. McIntyre and Dingman (55), concerned with transfer, found that low IQ subjects showed significant transfer of training when dealing with arithmetic problems beyond their MA level. Silverstein, and others (78) concluded that retarded subjects responded as normals to indefinite number terms such as "few," "some," and "a lot," indicating common basic understanding of quantity.

A very popular research problem in arithmetic learning is centered around the ability to conceive reversals (ability to learn to choose a stimulus which was previously incorrect). Heal, and others (35) concluded that compared to normals of the same MA, institutionalized retardates have (1) an inability to inhibit a previously acquired habit, (2) a susceptibility to disruption by novel stimuli, (3) a significant inferiority to conceive reversals.

Baumeister (3) studying recall abilities found normals, matched with retardates on chronological age (CA) only, significantly better on immediate recall, both incidental and intentional, and the retarded were superior on 48-hour retention of intentional material. He suggested that the learning deficit is task specific. One is reminded of Ellis' work on stimulus trace (29) along this same subject. Hetherington (38) in a CA-MA matched group found that familial and organic retarded did not differ in intentional learning or in retention. Neufeldt (64) concluded that short-term memory was an important difference between normal and retardates and postulated that the deficit was probably enhanced by retardates' lack of flexibility in search for and use of appropriate recall strategies. He further stated that the capacities of his organic and familial retardates were the same, but both were below MA normals in ability to adopt flexible modes of behavior and utilize ambiguous strategies of recall. Studies in learning research have the potential to be extremely useful in arithmetic if they can be translated into useful educational techniques for teaching.

Motivation Research

The results of arithmetic research in mental retardation, as one can observe, are replete with contradictory data. Nowhere is this more evident than in the motivation research. McPherson (57) points out that the reasons for conflicting results are due to the fact that various studies use different motivating techniques, study different aspects of learning, use different methods of sampling, and use subjects of differing intellectual levels and etiologies.

Schillo (76) concluded that the underachiever's expectation of negative reinforcement can affect his learning efficiency but that the expectancy function may be subject to several variables such as the difficulty level of the task, the presence of correction-reinforcement, and the ego-involving meanings which the task has for the subject. Although this study was done on reading, it seems to have relevance for arithmetic also. Anderson (2) noted that male achievers scored highest on arithmetic subtests of the Wechsler Intelligence Scale for Children (WISC) and confirmed the common finding that EMR girls achieved more in line with their ability than did EMR boys. Huber (40) found that one of his institutionalized high anxiety groups obtained significantly higher mean scores on reading, spelling, and mathematics than the low anxiety group. Klausmeier and Feldhusen (49) found significantly positive correlations between anxiety and arithmetic achievement in low IQ groups of retardates in nine special classes. They also found a significantly higher mean anxiety score for low IQ children than for average or high IQ children.

Dobbs (26) using Haywood's Motivator Hygiene Scale found that the "extrinsically" motivated EMR's performed better on her linear programmed arithmetic materials than the "intrinsically" motivated. This contradiction to expectation was explained by the dull, repetitious nature of the linear program which probably did not appeal to the intrinsically motivated child. Niedelman (65) concluded that in his EMR group, high achievers had high aspirations and low achievers had experienced prior failures and, therefore, had low aspirations. The reader is directed to Ellis (29) and Stevens and Heber (82) for motivational research in topics other than arithmetic abilities of the EMR.

Special Arithmetic Programs

The search for special teaching techniques and arithmetic programs is as old as mathematics itself. Research, however, on the effectiveness of programs is much more recent. Costello (18) concluded that math programs for the retarded must be concrete and preferably of the social utility type. Lehtinen and Strauss (50) placed emphasis on the abacus as a tool in a non-stimulation environment. Educators such as Kirk and Johnson (47) have suggested emphasis on simple counting, ability to recognize simple groups of numbers, concrete examples, and socially useful mathematics. In general, the efficacy of any of these methods is quite subjective and unsupported by empirical research.

In the 1960's a few research studies have been reported on specific arithmetic techniques with retarded children. Although they are not conclusive, and in some cases poorly done, they are steps in the right direction. Smith and Quackenbush (80) using Devereux teaching machines with EMR children compared arithmetic gains of twenty-three Ss with gains in other academic areas and with gains of similar students the previous year. They concluded that the teaching aids were useful in promoting academic achievement and motivation. Jacobs, and others (43) in a well-designed experiment using television, concluded that televised and conventional instruction were equally effective in imparting computational skills in five seventh-grade classes to pupils initially below norms in achievement. Two significant differences favored television and three nonsignificant differences were found. They also concluded that homogeneous grouping is more effective for television instruction. The authors did not mention mental retardation, and it is assumed that the subjects were just underachievers.

In a study with Sterns Structural Arithmetic, Ives (41) concluded that backward children appeared to learn arithmetic satisfactorily when taught by the "meaning method" regardless of apparatus, although the Stern apparatus group was superior in both quantity and quality of learning on a test designed for this study by Ives. Miller (61) in his master's thesis, investigating with Sterns equipment, concluded that the structural approach was a most effective aid to learning arithmetic for the retarded.

Eifermann and Etzian (28) in an ingenious experiment created a math system isomorphic to the number system 1–10, using colored rods to bypass the problem of previous learning. They concluded that a teaching method which emphasizes the reversible nature of math addition may enable a more comprehensive and flexible approach to arithmetic. Unfortunately, this was done with normal adults, but, as they suggest, could be a topic of research for EMR's.

Allen, and others (1) using Frostig-Horne materials with a small sample of retardates found significant increases in the experimental group in a combined score of figure ground discrimination, figure constancy, and spatial relations but not eye motor coordination and positioning in space. An interesting (though purely descriptive) study with Cuisinaire Rods presented by Callahan and Jacobson (12) showed that their EMR's could benefit from this type of arithmetic.

Rainey and Francis (72) made an assessment of the relative effectiveness of a commercial arithmetic program with teaching machines and two teacher-made programs. Using eight groups of intermediate EMR children, randomly assigned to treatment procedures, they found that the teacher-made program using rote methods provided significantly greater gains than either the commercial program or the teacher-made program using "understanding" principles. This study also found that there was a significant reading-arithmetic achievement interaction indicating that if the subject was reading above 2.3 grade placement he benefited more from the commercial program. Rainey and Francis suggest that early training be geared to rote methods which allow overlearning of basic facts.

Montague (62) concluded that the best technique is a good teacher when dealing with low socioeconomic children. Tate (84) programmed addition facts, 1–10, and found that retardates perform better using visual modality rather than the audio-visual system. McIntyre (56) concluded that repetitive items were fruitful in programmed arithmetic with retardates. Empirical research continues to be scanty on the effectiveness of special techniques for teaching arithmetic and perhaps holds the greatest promise in programmed instruction.

Programmed Instruction

Several methods of programmed instruction (PI) have been described by Stolurow (81), but the three major ones are those of Pressey (69), Skinner (79), and Crowder (20). The techniques used by these three have notable variations; step size, sequence of presentation, response type, and error rate. The effects of these differences have been studied at various levels of mental ability. Several researchers have reported little or no relationship between IQ and learning efficiency while some have found a significant relationship. Shay (77) reported no relationship between IQ and size of item step in a program dealing with Roman numerals.

Cartwright (16) did a study using mentally retarded adolescent subjects which showed that learning on unsystematic program presentation correlated most highly with different types of learning skills. Levin (51), however, randomly changed the difficulty order of items and found no effect on acquisition, retention, or transfer of learning. Some researchers have shown branching programs to be more effective with higher ability pupils than with low ability pupils (13, 19, 73).

Fry (32) reported constructed responses to be superior to multiple-choice responses in high school students, as measured by posttest scores. There was, however, a tendency for students to take longer to complete constructed response programs. Briggs (9) found multiple-choice responses to be superior to constructed responses. Price (70) using addition and subtraction PI with retardates in a residential setting, reported multiple-choice responses to be superior in subtraction. There was no difference in addition.

Little or no correlatioin was found between IQ and learning efficiency when error rates were low during the administration of PI (25, 37, 60, 68). This finding was strengthened by the results of a study by Keisler (45), who reported a positive relationship between mental age and arithmetic gain scores using an arithmetic program with a high error rate.

Working with mental retardates, McIntyre and Dingman (55) used repetition of items and showed decreasing error rates with increasingly difficult items. A similar finding was reported by Weir and Stevenson (88). McIntyre and Dingman concluded that PI based on simple conditioning was more effective with mental retardates. Later in a post-hoc analysis of their data, McIntyre and Dingman (55) concluded: (1) When programmed material is of low difficulty, the type of learning curve is the same for different learning abilities, but the lower IQ takes longer. (2) When programs get more difficult, the Ss with better experience respond better. (3) The fatigue factor is important. Malpass, and others (53) used massed repetitions in word recognition programming with mental retardates and found that it apparently facilitated learning in the early stages. McIntyre (56), however, reported no differences between massed and distributed practice.

The results of a study by Porter (68) seem to indicate that programming a course of study may actually lower the relationship of IQ and classroom learning. No correlation was found between the IQ and achievement of elementary school pupils using PI in spelling, while a significant correlation was found between the IQ and achievement of controls in a regular classroom spelling program. In support of Porter's results, Price's (70) study showed that only eighty-six class periods using PI, as compared with 130 class periods using regular classroom teaching methods, were required to learn the same amount of material. Price stated that all three of his retarded groups showed significant improvement in addition, but only the group on Multiple Choice PI had significant improvement in subtraction. Merachnik (58), however, found no difference between regular methods of teaching and PI in addition, subtraction, multiplication, and division in a pilot project with educable mental retardates. In a study by Northcutt (66) the learning of pupils in the regular classroom instructional program was superior to that of pupils using PI.

Briggs and Angell (10) reviewed nineteen studies in which the effectiveness of PI was compared with conventional teaching procedures in science and mathematics. "No significant differences" were reported more often than significant differences. In a 3-year study by Blackman, and others (7) there was no significant difference between the arithmetic achievement of pupils using PI and those continuing in the regular classroom instructional program, as measured by the Wide Range Achievement Test and the Metropolitan Achievement Test. Using a measure designed specifically for the PI content, however, there was a significant difference in the arithmetic achievement of pupils under the two methods of instruction. Operationally defined standards for PI seem necessary before conflicting evidence can be resolved. Herrick (37) states that, in her estimation, good programs are still to be written, and that failure thus far has been due to an inability to produce effective programs rather than to a failure of the method. The particular program used seems to be the crucial factor in the outcome of research on programmed instruction.

Dobbs (26), in a 60-day treatment, used a linear Arithmetic Program designed to eliminate reading as a variable and fitted to the achievement level of her subjects. She found that her thirteen randomly selected EMR experimental groups made significant gains over thirteen EMR control groups who were given regular classroom methods of arithmetic instruction.

The literature reviewed does not provide the desired guidance for locating or selecting PI most appropriate to specific mental ability groups. In general, it seems that step size (77) and low error rates (25, 36, 60, 68) decrease the relationship between IQ and learning efficiency; repetition provides for greater learning efficiency in mental retardates (88, 55); and branching programs appear best for higher ability levels (13, 19, 73).

Summary and Conclusions

An attempt has been made to review the literature dealing with teaching arithmetic to mentally retarded children. There are many conclusions formulated throughout this paper, but a brief summary of them must include these major findings:

1. Cruickshank's 1948 description of the arithmetic characteristics of EMR's is still valid.
2. Dunn's conclusion that retardates tend to perform more closely to their MA in computation than arithmetic reasoning is still accurate although there is growing evidence that retardates can improve reasoning abilities (skills).
3. Arithmetic concepts are formulated by maturation stages, and although a retardate can be drilled to perform computation, it does not necessarily imply an understanding.
4. The endogenous-exogenous etiology theory has little utility for arithmetic.
5. Research in learning has much to offer if it can be communicated to teachers.
6. Special teaching methods, though high in face validity, have almost no empirical research support.
7. Programmed instruction can be a useful method to teach arithmetic to retardates.

REFERENCES

1. Allen, R. M.; Deckman, I.; Haupt, T. D, "A Pilot Study of the Immediate Effectiveness of the Frostig-Horne Training Program with Educable Retardates," *Exceptional Children,* 33:(1)41–41, 1966.
2. Anderson, L. M., "Factors Affecting High and Low Achievement Among Adolescents Enrolled in Special Classes for Mentally Handicapped," *Dissertation Abstracts,* 25: 279, 1965.
3. Baumeister, A., "Comparison of Normals and Retardates with Respect to Incidental and Intentional Learning," *American Journal of Mental Deficiency,* 68: 404–408, 1963.
4. Bennett, A., *A Comparative Study of Subnormal Children in the Elementary Grades,* Teachers College, Columbia University, New York, 1932.
5. Bensburg, G. L., "The Relation of Academic Achievement of Mental Defectives to Mental Age, Sex, Institutionalization and Etiology," *American Journal of Mental Deficiency,* 58:327–330, 1953.
6. Benton, A.; Hutcheon, J.; Seymour, E., "Arithmetic Ability, Finger-localization Capacity and Right–left Discrimination in Normal and Defective Children," *American Journal of Orthopsychiatry,* 21:756–766, 1956.
7. Blackman, L. S.; Smith, M. P., "*The Developement and Evaluation of a Curriculum for Educable Mental Retardates Utilizing Self-instructional Devices or Teaching Machines,*" Title VII, Project No. 368, Edward R. Johnstone Training and Research Center, Bordentown, New Jersey, 1964.
8. Blatt, B., "The Physical, Personality, and Academic Status of Children Who are Mentally Retarded Attending Special Classes as Compared with Children who are Mentally Retarded Attending Regular Classes," *American Journal of Mental Deficiency,* 62:810–818, 1958.
9. Briggs, L. J., "The Effect of Response Mode and Learning Material Difficulty on Automated Instruction," in *Teaching-Machine Conference Proceedings,* University of Southern California, Los Angeles, 1960, pp. 19–20.
10. Briggs, L. J.; Angell, D., "Programmed Instruction in Science and Mathematics," *Review of Educational Research,* 34: 354–373, 1964.
11. Brown, I.; Dyer, L., "Social Arithmetic Training for the Subnormal: A Comparison of Two Methods,"

Journal of Mental Subnormality, 9: 8–12, 1963.

12. Callahan, J.; Jacobson, R., "An Experiment with Retarded Children and Cuisenaire Rods," *Arithmetic Teacher*, 14: (1) 19–13, 1967.
13. Campbell, V. N., *Studies of Bypassing as a Way of Adapting Self-instructional Programs to Individual Differences*, Title VII, Grant No. 736103, U. S. Office of Education, 1962.
14. Capobianco, R., "Quantitative and Qualitative Analyses of Endogenous and Exogenous Boys on Arithmetic Achievement," *Monograph of Social Research in Child Development*, 19, 1954.
15. Carrison, D.; Werner, H., "Principles and Methods of Teaching Arithmetic to Mentally Retarded Children," *American Journal of Mental Deficiency*, 47: 309–317, 1943.
16. Cartwright, G. P., "Two Types of Programmed Instruction for Mentally Retarded Adolescents," unpublished master's theses, University of Illinois, 1962.
17. Chase, C., "Application of Levels of Concept Formation to Measurement of Vocabulary," *Journal of Educational Research*, 55: 75–78, 1961.
18. Costello, H., "The Responses of Mentally Retarded Children to Specialized Learning Experiences in Arithmetic," unpublished doctoral dissertation, University of Pennsylvania, 1941.
19. Coulson, J.; Silverman, H., "Effects of Three Variables in a Teaching Machine," *Journal of Educational Psychology*, 51: 135–143, 1960.
20. Crowder, N., "Automatic Tutoring by Intrinsic Programming," in Lunsdaine, A.; Glaser, R., (eds.), *Teaching Machines and Programmed Learning*, National Educational Association, Washington, D. C., 1960, pp. 286–298.
21. Cruickshank, W. M., "Arithmetic Ability of Mentally Retarded Children: I. Ability to Differentiate Extraneous Materials from Needed Arithmetic Facts," *Journal of Educational Research*, 42: 161–170, 1948, (a).
22. Cruickshank, W. M., "Arithmetic Ability of Mentally Retarded Children: II. Understanding Arithmetic Processes," *Journal of Educational Research*, 42: 279–288, 1948, (b).
23. Cruickshank, W. M., "Arithmetic Work Habits of Mentally Retarded Boys," *American Journal of Mental Deficiency*, 52: 318–330, 1948, (c).
24. Denny, M. R., "Learning," in Stevens, H.; Heber, R., (eds.), *Mental Retardation: A Review of Research*, University of Chicago Press, Chicago, 1964.
25. Detambel, M. H.; Stolurow, L. M., "Stimulus Sequence and Concept Learning," *Journal of Experimental Psychology*, 51: 34–40, 1956.
26. Dobbs, V. H., *Motivational Orientation and Programmed Instruction Achievement Gain of Educable Mentally Retarded Adolescents*, (Doctoral dissertation, George Peabody College), University Microfilms, Ann Arbor, Michigan, 1967. No. 68–2893.
27. Dunn, L. M., "A Comparison of the Reading Processes of Mentally Retarded Boys of the Same Mental Age," *Monographs of the Society for Research in Child Development*, 19: (1), 1954.
28. Eifermann, R.; Etzion, D., "Awareness of Reversibility: Its Effect on Performance of Converse Arithmetic Operations," *British Journal of Educational Psychology*, 34: (2) 151–157, 1964.
29. Ellis, N. R., (ed.), *Handbook of Mental Deficiency*, McGraw-Hill, New York, 1963.
30. Estes, B.; Combs, A., "Perception of Quantity," *Journal of Genetic Psychology*, 108: (2) 333–336, 1966.
31. Finley, C., "Arithmetic Achievement in Mentally Retarded Children: The Effects of Presentation of the Problem in Different Contexts," *American Journal of Mental Deficiency*, 67: 281–286, 1963.
32. Fry, E., "Teaching Machines: An Investigation of Constructed Versus Multiple-choice Methods of Response," unpublished doctoral dissertation, University of Southern California, 1960.
33. Gallagher, J., "A Comparison of Brain-injured and Non-brain-injured Mentally Retarded Children on Several Psychological Variables," *Monographs of the Society for Research in Child Development*, 22: (2), 1957.
34. Goldstein, H.; Moss, J.; Jordan, L., *The Efficacy of Special Class Training on the Development of Mentally Retarded Children*, Cooperative Research Project No. 619, University of Illinois, 1965.
35. Heal, L.; Ross, L. E.; Sanders, B., "Reversal and Partial Reversal in Mental Defectives and Normal Children of a Comparable Mental Age," *American Journal of Mental Deficiency*, 71: 411–416, 1966.
36. Herrick, M. C., *The Effect of Problem-setting Questions on Rate and Amount of Learning in Programming Teaching Machines*, Title VII, Grant No. 712130, U. S. Office of Education, 1962.
37. Herrick, M. C., "Innovations of Education-programmed Instruction," *School Science and Mathematics*, 65: 405–408, 1965.
38. Hetherington, E.; Banta, T., "Incidental and Intentional Learning in Normal and Mentally Retarded Children," *Journal of Comparative and Physiological Psychology*, 55: 402–404, 1962.
39. Hood, H., "An Experimental Study of Piaget's Theory of Development of Numbers in Children," *British Journal of Psychology*, 53: 273–286, 1962.
40. Huber, W., "The Relationship of Anxiety to Academic Performance of Institutionalized Retardates," *American Journal of Mental Deficiency*, 69: (4) 462–466, 1965.
41. Ives, L., "Value of the Stern Arithmetic Apparatus in a Backward Class," *British Journal of Educational Psychology*, 32: 300–303, 1962.
42. Jacobs, J. N., "A Study of Performance of Slow Learners in the Cincinnati Public Schools on Mental and Achievement Tests," *American Journal of Mental Deficiency*, 63: 238–243, 1957.
43. Jacobs, J.; Bollenbacher, J.; Keiffer, M., "Teaching 7th Grade Math by Television to Homogenously Grouped Below Average Pupils," *Mathematics Teacher*, 59: 551–555, 1961.
44. Jarvis, O., "Time Allotment Relationships to Pupil Achievement in Arithmetic," *Arithmetic Teacher*, 10: 248–250, 1963.
45. Keisler, E., "The Development of Understanding in Arithmetic by a Teaching Machine," *Journal of Educational Psychology*, 50: 247–253, 1959.
46. Kirk, S. A., "Research in Education," in Stevens, H.; Heber, R. (eds.), *Mental Retardation: A Review of Research*, University of Chicago Press, Chicago, 1964.
47. Kirk, S.; Johnson, O., *Educating the Retarded Child*, Houghton Mifflin Company, Cambridge, Massachusetts, 1951.
48. Klausmeier, H.; Feldhusen, J.; Check, J., *An Analysis of Learning Efficiency in Arithmetic of Mentally Retarded Children in Comparison with Children of Average and High Intelligence*, Project No. 153, U. S. Office of Education, 1959.
49. Klausmeier, H.; Feldhusen, J., "Anxiety, Intelligence and Achievement in Children of Low, Average, and High Intelligence," *Child Development*, 33: 403–409, 1962.
50. Lehtinen, L.; Strauss, A., "Arithmetic Fundamentals for the Brain-crippled Child," *American Journal of Mental Deficiency*, 49: 149–154, 1944.
51. Levin, G., *Principles of Programming Material for Teaching Machines and Their Relation to Transfer of Training*, Title VII, Project 538, U. S. Office of Education, 1961.
52. Linhart, I., Uchenie i undstevennoe razvitei, (Learning and Mental Development), *Voprosy Psikhologu*, 4: 32–38, 1966, presented at the 1966 International Psychological Congress in Moscow.
53. Malpass, L.; Gilmore, A.; Mardy, M.; Williams, C.,

“Comparison of Two Automated Teaching Procedures for Retarded Children: A Progress Report,” unpublished manuscript, MacDonald Training Center, Tampa, Florida, 1962.

54. Matthews, C.; Folk, E., “Finger Localization, Intelligence, and Arithmetic in Mentally Retarded Subjects,” *American Journal of Mental Deficiency*, 69: (1), 107–113, 1964.

55. McIntyre, R.; Dingman, H., “Mental Age Vs. Learning Ability: An Investigation of Transfer of Training Between Hierarchial Levels,” *American Journal of Mental Deficiency*, 68: 396–403, 1963.

56. McIntyre, R. B., “Effects of Repetitious Programming in the Acquisition of Addition Facts by Educable Retardates,” unpublished doctoral dissertation, George Peabody College, 1964.

57. McPherson, M., “Learning and Mental Deficiency,” *American Journal of Mental Deficiency*, 62: 870–877, 1958.

58. Merachnik, D., *Small Sample Pilot Project for the Acquisition and Refinement of Arithmetic Computation Skills with the Mentally Retarded, Via Programmed Instruction*, Union County Regional High School, District No. 1, Berkeley Heights, New Jersey, 1962.

59. Merrill, M. A., “The Relation of Intelligence to Ability in the Three R’s in the Case of Retarded Children,” *Journal of Genetic Psychology*, 28: 249–274, 1921.

60. Meyer, S., “Report on the Initial Test of a Junior High School Vocabulary Program,” in Lunsdaine, A.; Glaser, R. (eds.), *Teaching Machines and Programmed Learning*, National Education Association, Washington, D. C., 1960, pp. 229–245.

61. Miller, J., “Investigating the Structural Arithmetic Program of Catherine Stern,” unpublished master’s thesis, Boston University School of Education, Massachusetts, 1960.

62. Montague, D., “Arithmetic Concepts of Kindergarten Children in Contrasting Socio–economic Areas,” *Elementary School Journal*, 64: 393–397, 1964.

63. Mullen, F.; Itkin, W., “Achievement and Adjustment of Educable Mentally Handicapped Children,” Cooperative Research Project (6529), Board of Education City of Chicago, 1961.

64. Neufeldt, A., “Short-term Memory in the Mentally Retarded: An Application of the Dichotic Listening Technique,” *Psychological Monographs*, 80:(12)1–31, 1966.

65. Niedelman, J., “The Relationship Between Level of Aspiration and School Achievement in the Educable Mentally Retarded Adolescent,” *Dissertation Abstracts*, 26: (8) 4814, 1966.

66. Northcutt, M., “The Comparative Effectiveness of Classroom and Programmed Instruction in the Teaching of Decimals to Fifth Grade Students,” unpublished doctoral dissertation, George Peabody College, 1963.

67. Piaget, J., “How Children Form Mathematical Concepts,” *Scientific American*, November, 1953.

68. Porter, D., “Some Effects of Year Long Teaching Machine Instruction,” Galanter, E. H. (ed.), *Automatic Teaching: The State of the Art*, Wiley, New York, 1959, pp. 187–198.

69. Pressy, S., “Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-instruction,” *Journal of Psychology*, 29: 417–447, 1950.

70. Price, J., “Automated Teaching Programs with Mentally Retarded Students,” *American Journal of Mental Deficiency*, 63: 69–72, 1963.

71. Quay, L. C., “Academic Skills,” in Ellis, N. R. (ed.), *Handbook of Mental Deficiency*, McGraw-Hill Book Co., New York, 1963.

72. Rainey, D.; Francis, K., “An Evaluation of a Programmed Textbook with Educable Mentally Retarded Children,” *Exceptional Children*, 34: (3) 1969–174, 1967.

73. Rigney, J., “Experimental Evaluation of Autoinstructional Concepts in Navy Training Environments,” *Teaching-machine Conference Proceedings*, University of Southern California, Los Angeles, 1960.

74. Ring, S. B., “A Comparison of Achievement and MA of Ninety-eight Special Class Children,” unpublished master’s thesis, Boston University, 1951.

75. Schilder, P., *The Image and Appearance of the Human Body*, International University Press, New York, 1935.

76. Schillo, R., “Concept Learning of Achievers and Underachievers as a Function of Task Expectancy,” *Dissertation Abstracts*, 24:(9)3841, 1964.

77. Shay, C., “The Relationship of Intelligence to Size of Item Step on a Teaching Machine Program,” unpublished doctoral dissertation, University of California at Los Angeles, 1960.

78. Silverstein, A.; Auger, R.; Krudis, B., “The Meaning of Indefinite Number Terms for Mentally Retarded Children,” *American Journal of Mental Deficiency*, 69: 419–429, 1964.

79. Skinner, B., “Teaching Machines,” *Scientific American*, vol. 128, 1959.

80. Smith, A.; Quackenbush, J., “Devereux Teaching Aids Employed in Presenting Elementary Mathematics in a Special Education Setting, *Psychological Review*, 7: 333–336, 1960.

81. Stolurow, L. M., “Teaching Machines and Special Education,” *Educational and Psychological Measurement*, 20: 429–448, 1960.

82. Stevens, H.; Heber, R., *Mental Retardation*, University of Chicago Press, Chicago, 1964.

83. Strauss, A.; Werner, H., “Deficiency in the Finger Schema in Relation to Arithmetic Disability (Finger agnosia and acalculia), *American Journal of Orthopsychiatry*, 8: 719–725, 1938.

84. Tate, B., “Effectiveness of Retardate Learning of Arithmetic as a Function of Mode of Stimulus Presentation by an Auto–instructional Devise,” *Dissertation Abstracts*, 25: 6401, 1964.

85. Thurstone, T. G., *An Evaluation of Educating Mentally Handicapped Children in Special Classes and in Regular Classes*, U. S. Office of Education Cooperative Residence Project, Project OE–SAE #6452, School of Education, University of North Carolina, Chapel Hill, 1959.

86. Voerg, M., “A Comparison of Learning Between Mentally Deficient, Mentally Retarded, and Normals,” *Dissertation Abstracts*, 25: 3714, 1964.

87. Wang, H.; Liu, C.; Fen, T., “An Investigation into the Development of Concepts in Children 4–9 years,” *Acta Psychologia Sinica*, 4: 352–360, 1964.

88. Weir, M.; Stevenson, H., “The Effect of Verbalization in Children’s Learning as a Function of Chronological Age,” *Child Development*, 30: 143–149, 1959.

89. Werner, H.; Carrison, D., “Measurement and Development of the Finger Schema in Mentally Retarded Children: Relation of Arithmetic Achievement to Performance on the Finger Schema Test,” *Journal of Educational Psychology*, 33: 252–264, 1942.

90. Wilson, F. T., “Some Achievements of Pupils of the Same Mental Ages but Different Intelligence Quotients,” *Journal of Educational Research*, 14: 43–53, 1926.

91. Woodward, M., “Concepts of Number of the Mentally Subnormal Studied by Piaget’s Method,” *Journal of Child Psychology and Psychiatry*, 2: 249–259, 1961.

92. Young, M., “Academic Requirements of Jobs Held by Educable Mentally Retarded in State of Connecticut,” *American Journal of Mental Deficiency* 62: 792, 802, 1956.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 4, December 1970)

Characteristics of Students Who Could Not Take and Schools Which Did Not Offer Driver Training[1]

JOE E. SHIVELY and WILLIAM ASHER, Purdue University

ABSTRACT

In this study essentially a random sample of 798 U. S. high school seniors' data was selected from a data bank. Scores were available on seventy variables related to driving safety and a criterion variable, whether they had taken driver training or not taken it because it was not offered. Eighteen of these variables showed differences between the two groups. These variables cluster around academic knowledge, intelligence, and socioeconomic status. The 861 U. S. high schools' data were also selected from a data bank. Scores were available on ninety-seven variables related to driving safety and a criterion variable, whether or not they offered a driver training course. Forty-eight of these variables showed differences between the two groups. These variables cluster around socioeconomic status. It is concluded that different kinds of students do and do not take driver training courses. Schools which do and do not offer driver training are also different on variables which also relate to frequency of automobile accidents. While this research does not test the hypothesis that driver training courses make no difference in number of automobile accidents, it does demonstrate further that variables other than taking or not taking and offering or not offering driver training differentiates between these groups.

THERE IS currently a national mandate on instituting driver training courses into school systems. Public Law 89–10 provided funds for improving driver training, and Public Law 89–564 required states to provide driver training programs in the school systems or be subject to a 10 percent cut in Federal highway construction subsidies (12).

Often driver training is justified on the basis of accident reduction. However, Asher (1) has found variables other than taking or not taking driver training, *even though available*, to differentiate between these groups. McGuire and Kersh (10) also indicate variables other than the variable of taking or not taking driver training to be related to the ultimate criterion variable of highway accident frequency, accident severity, or violation frequency. However, insurance companies often reduce rates for individuals who have had driver training because it is known that students who *do not* take driver training have more accidents (12). But most studies which indicate that driver training reduces accidents on a cause and effect basis are open to question scientifically.

Analysis of data involving characteristics of students who took driver training and students who were *unable* to take driver training because it was not offered in their schools might yield important clues as to the differences in accident rates of the two groups of students. In addition, since it is the school which offers driver training and not the individual student, analysis of data involving school characteristics in relationship to driver training might well yield important clues as to the differences in accident rates of the two groups of students. It might also help to explain further why accident rates seem to be lower for students taking driver training.

Thus the questions are asked: Do those students who take and cannot take driver training (because it is not offered in their schools) differ on variables which are related to accident related variables? Do those schools which offer and do not offer driver training differ on variables which are related to accident related variables?

Studies on sociological and psychological predictor variables of automobile accidents are cited in Goldstein (8) which indicate that years of education, general intelligence, and general aptitude have a negative relationship with accidents. Asher (1) indicates that those variables relating to the decision to take driver training cluster around academic knowledge, intelligence, and socioeconomic status. Conger, Miller, and Rainey (4) in a study of accident and violation records found that significant differences exist between the high and low accident and violation groups on exposure (miles driven per year), socioeconomic status, and IQ.

Since driver training is costly, those schools which do not offer it may be in lower socioeconomic communities. Hence socioeconomic related variables might well differentiate between those *schools* which offer driver training and those which do not offer it. Similarly, socioeconomic related variables would differentiate between those *students* who take driver training and those who do not take driver training because it is not offered in their schools. Since socioeconomic variables are related to automobile accidents, a further study of students unable to take and schools not offering driver training, could lead to an improved understanding of the relationship between driver training and automobile accidents.

Method

Project TALENT is a research project with major support from the U. S. Office of Education which collected many types of data from about 400,000 students. School and community data were also obtained on a 5 percent sample of all the U. S. high schools. The initial results are reported in *The American High School Student* (5) and *Studies of the American High School* (6). The data are available in a data bank.

From the Project TALENT data bank certain characteristics of a sample of 798 students were selected for study. Seventy variables were chosen in addition to the criterion variable of having taken driver training or not being able to take it because it was not offered. The seventy variables were the same as used in Asher (1). Part of question 296 was used as the criterion for groupings. It read, "Did you take driver training at school?" The alternatives were A, "Yes" and B, "No, this course is not available." (Asher (1) has analyzed data for students who answered alternative C, "No, although this course is available." Alternative D, "No, but I expect to take the course in a later grade," has no meaning for this sample of seniors since there was no way that they could expect to take the course in a later grade.)

From the Project TALENT data bank certain characteristics of the sample of schools were also selected for study. Forty-seven items from the General School Characteristics questionnaire (7) were chosen in addition to the criterion variable for the sample of 861 high schools: offering driver education or not offering driver education. The forty-seven items were selected for their possible relevancy to socioeconomic studies of the community. Such general items as class size; absenteeism; number and educational level of teaching staff; number of students per grade; percent male and female per grade; school expenditures; racial percentages; delinquency and dropout percentages; characteristics of community housing, property evaluation, and per pupil expenditures; and the availability of an automobile mechanic course were included.

Specifically they were the following variables as detailed in the GSC questionnaire (7) : *test item 6*, number of school days; *9 and 10*, class size in science and non-science courses; *13*, special classes; *16*, advanced placement opportunities; *29*, percent daily absenteeism; *32, 33, and 34*, library facilities and physical plant characteristics; *36*, school type (public versus non-public); *43, 47, 48, and 51*, number of teachers and their qualifications; *56 and 58*, salary and experience of teachers; *59 and 60*, sex and transient percentages; *62 through 67*, school programs percentages; *69 through 74*, delinquency dropout and advanced training percentages; *82*, percentage on half-day session; *84*, percent of PTA membership; *86 through 92*, characteristics of community housing, property evaluation, and per pupil expenditures; *93 and 94*, familial language and school support; *95 through 98*, nationality percentages; *100*, community facilities; and the availability of the automobile mechanics course as cited under the category, Trade and Industrial Education-Vocational. The criterion variable of availability of driver education, listed under the category Health, Safety, and Physical Education, was also selected.

The questionnaire, which contains 108-item variables and approximately 375 course listings, was filled out by the principal (or, if not available, the assistant principal or dean) of the school.

Analysis of the data involved using the missing data BIMED program (03D) (3) for generating point-biserial, phi, and other product-moment correlation coefficients to determine which variables will differentiate between driver training and no driver training groups.

These correlations or their functions can be tested for a significant difference from zero. For the school data a phi must reach .067 ($n = 861$) to be significant at the .05 level, a χ^2 of 3.865. A phi of .088 ($n = 861$) to be significant at the .01 level, a x^2 of 6.668. For the point–biserial correlations, rhos, and standard correlations Guilford's Table D (9) was used and significance is obtained when the coefficients reach a value of .067 at the .05 level ($n = 861$) or .088 at the .01 level ($n = 861$). (The values were obtained from a linear interpolation of the coefficients presented for $n = 500$ and $n = 1000$.) For the student data a phi must reach .098 ($n = 400$) to be significant at the .05 level or .128 ($n = 400$) at the .01 level. Guilford's Table D (9) was again used for point-biserials, rhos, and standard correlations, and significance is obtained when the coefficients reach a value of .098 at the .05 level ($n = 400$) or .128 at the .01 level ($n=400$). (For the variable "knowledge of mechanics," significance was declared at the .05 level when the coefficient reached a value of .092 ($n = 451$).)

Results

Students. From the 1960 sample of 798 U. S. high school seniors, 303 took driver training and 151 could not take driver training because it was not available. (The remaining 344 seniors are those who answered alternatives "C" or "D" of SIB variable 296.)

For the sample of students, of the correlations between the seventy variables and taking driver training, eighteen were significant at the .05 level or beyond. Specifically, twelve were significant at the .01 level and six were significant at the .05 level. (The expected number of significant relationships by chance alone would be 3.5 if all of the seventy variables were independent of each other.)

The variables that are significant at the .01 level are knowledge of music, social studies, physical sciences, biological sciences, scientific attitude, and electricity and electronics; mechanical reasoning, preference as a speed test of choice, interest in computation, IQ composite, socioeconomic index, and the turning in of sloppy assignments.

The variables that are significant at the .05 level are knowledge of literature, mathematics, sports, and mechanics; number of dates per week, the student's self-report of mathematics grades. The eighteen significant correlation coefficients are presented in Table 1.

Table 1.—Correlation of Selected TALENT Student Variables with SIB Variable 296

Project TALENT Variable	Correlation Coefficient
Knowledge of:	
Literature	-0.109 *
Music	-0.207 **
Social Studies	-0.149 **
Mathematics	-0.108 *
Physical Sciences	-0.146 **
Biological Sciences	-0.134 **
Scientific Attitude	-0.139 **
Electricity and Electronics	-0.103 **
Sports	-0.122 *
Mechanics	-0.093 *
Mechanical Reasoning	-0.155 **
Preferences:	
Speed of Choice	0.150 **
Interests:	
Computation	0.136 **
I.Q. Composite	-0.145 **
Socio-economic Index	-0.160 **
Social Activities:	
Dates per Week	0.121 *
Study Habits:	
Sloppy Assignments	-0.130 **
Self-report:	
Mathematics Grades	-0.109 *

*p < .05 **p < .01

NOTE: SIB Variable 296 had the following numeric code:

Took Driver Training = 1
Did Not Take Driver Training (Not offered in their school) = 2

With those Not Taking Driver Training having been assigned the higher numeric code value, a negative correlation between this variable and another indicates that non-takers received lower scores on the other variables. The signs of the correlation are in a sense an artifact of the coding system.

Schools. Of the 861 schools, 508 offered a driver training course, 59 percent. Of the ninety-seven relationships of prime interest in this study, forty-one were significant at the .01 level and seven were significant at the .05 level. (The expected number of significant relationships by chance alone would be 4.85 at the .05 level or .97 at the .01 level if all of the ninety-seven variables were independent of each other.)

The variables that are significant at the .01 level are number offering special class for speech difficulty, number of books in the library, type of school, number of full-time teachers, number of full-time teachers who left in past year, number of full-time teachers with a bachelors degree, percent full-time certified teachers, teacher's starting salary (BS, male), numbers of tenth-, eleventh-, and twelfth-grade male students, total number of males, numbers of tenth-, eleventh-, and twelfth-grade female students, total number of females, numbers of tenth-, eleventh-, and twelfth-grade males and females, total number of males and females in all grades, offering commercial or distributive diplomas, industrial, trade or vocational diplomas, and agricultural diplomas, number of tenth-grade male dropouts, number of tenth-grade female dropouts, number of twelfth-grade male dropouts, number of twelfth-grade female dropouts, percent on half-day sessions, type of residence in school area, percentage school support from local sources, percentage of American Indians, percentage of Negro pupils, community facilities available including public library, 4-H club, Boy Scouts, etc., Girl Scouts, etc., Future Nurses Club, Future Farmers of America, Future Teachers of America, and availability of an auto mechanics course.

The variables that are significant at the .05 level are offering special class for low IQ or mentally retarded, offering special class for mathematics difficulty, concerts, average age of pupils homes, community facilities available including community theater and public swimming facility, and region of U.S. in which school is located. The forty-eight significant correlation coefficients are presented in Table 2.

Discussion

The organization and patterning of the significant relationships fit into a general theory and relate to the prior research in such a way that these relationships appear to form meaningful results.

Interpretation of the results indicates that the variables which differentiated between schools

schools which offer driver training and those which do not offer it. Similarly, socioeconomic related variables would differentiate between those *students* who take driver training and those who do not take driver training because it is not offered in their schools. Since socioeconomic variables are related to automobile accidents, a further study of students unable to take and schools not offering driver training, could lead to an improved understanding of the relationship between driver training and automobile accidents.

Method

Project TALENT is a research project with major support from the U. S. Office of Education which collected many types of data from about 400,000 students. School and community data were also obtained on a 5 percent sample of all the U. S. high schools. The initial results are reported in *The American High School Student* (5) and *Studies of the American High School* (6). The data are available in a data bank.

From the Project TALENT data bank certain characteristics of a sample of 798 students were selected for study. Seventy variables were chosen in addition to the criterion variable of having taken driver training or not being able to take it because it was not offered. The seventy variables were the same as used in Asher (1). Part of question 296 was used as the criterion for groupings. It read, "Did you take driver training at school?" The alternatives were A, "Yes" and B, "No, this course is not available." (Asher (1) has analyzed data for students who answered alternative C, "No, although this course is available." Alternative D, "No, but I expect to take the course in a later grade," has no meaning for this sample of seniors since there was no way that they could expect to take the course in a later grade.)

From the Project TALENT data bank certain characteristics of the sample of schools were also selected for study. Forty-seven items from the General School Characteristics questionnaire (7) were chosen in addition to the criterion variable for the sample of 861 high schools: offering driver education or not offering driver education. The forty-seven items were selected for their possible relevancy to socioeconomic studies of the community. Such general items as class size; absenteeism; number and educational level of teaching staff; number of students per grade; percent male and female per grade; school expenditures; racial percentages; delinquency and dropout percentages; characteristics of community housing, property evaluation, and per pupil expenditures; and the availability of an automobile mechanic course were included.

Specifically they were the following variables as detailed in the GSC questionnaire (7

number of school days; *9 and 10,* class size in science and non-science courses; *13,* special classes; *16,* advanced placement opportunities; *29,* percent daily absenteeism; *32, 33, and 34,* library facilities and physical plant characteristics; *36,* school type (public versus non-public); *43, 47, 48, and 51,* number of teachers and their qualifications; *56 and 58,* salary and experience of teachers; *59 and 60,* sex and transient percentages; *62 through 67,* school programs percentages; *69 through 74,* delinquency dropout and advanced training percentages; *82,* percentage on half-day session; *84,* percent of PTA membership; *86 through 92,* characteristics of community housing, property evaluation, and per pupil expenditures; *93 and 94,* familial language and school support; *95 through 98,* nationality percentages; *100,* community facilities; and the availability of the automobile mechanics course as cited under the category, Trade and Industrial Education-Vocational. The criterion variable of availability of driver education, listed under the category Health, Safety, and Physical Education, was also selected.

The questionnaire, which contains 108-item variables and approximately 375 course listings, was filled out by the principal (or, if not available, the assistant principal or dean) of the school.

Analysis of the data involved using the missing data BIMED program (03D) (3) for generating point-biserial, phi, and other product-moment correlation coefficients to determine which variables will differentiate between driver training and no driver training groups.

These correlations or their functions can be tested for a significant difference from zero. For the school data a phi must reach .067 ($n = 861$) to be significant at the .05 level, a χ^2 of 3.865. A phi of .088 ($n = 861$) to be significant at the .01 level, a χ^2 of 6.668. For the point–biserial correlations, rhos, and standard correlations Guilford's Table D (9) was used and significance is obtained when the coefficients reach a value of .067 at the .05 level ($n = 861$) or .088 at the .01 level ($n = 861$). (The values were obtained from a linear interpolation of the coefficients presented for $n = 500$ and $n = 1000$.) For the student data a phi must reach .098 ($n = 400$) to be significant at the .05 level or .128 ($n = 400$) at the .01 level. Guilford's Table D (9) was again used for point-biserials, rhos, and standard correlations, and significance is obtained when the coefficients reach a value of .098 at the .05 level ($n = 400$) or .128 at the .01 level ($n=400$). (For the variable "knowledge of mechanics," significance was declared at the .05 level when the coefficient reached a value of .092 ($n = 451$).)

Results

Students. From the 1960 sample of 798 U. S. high school seniors, 303 took driver training and was not available. (The remaining 344 seniors are those who answered alternatives "C" or "D" of SIB variable 296.)

For the sample of students, of the correlations between the seventy variables and taking driver training, eighteen were significant at the .05 level or beyond. Specifically, twelve were significant at the .01 level and six were significant at the .05 level. (The expected number of significant relationships by chance alone would be 3.5 if all of the seventy variables were independent of each other.)

The variables that are significant at the .01 level are knowledge of music, social studies, physical sciences, biological sciences, scientific attitude, and electricity and electronics; mechanical reasoning, preference as a speed test of choice, interest in computation, IQ composite, socioeconomic index, and the turning in of sloppy assignments.

The variables that are significant at the .05 level are knowledge of literature, mathematics, sports, and mechanics; number of dates per week, the student's self-report of mathematics grades. The eighteen significant correlation coefficients are presented in Table 1.

Table 1.—Correlation of Selected TALENT Student Variables with SIB Variable 296

Project TALENT Variable	Correlation Coefficient
Knowledge of:	
Literature	-0.109 *
Music	-0.207 **
Social Studies	-0.149 **
Mathematics	-0.108 *
Physical Sciences	-0.146 **
Biological Sciences	-0.134 **
Scientific Attitude	-0.139 **
Electricity and Electronics	-0.103 **
Sports	-0.122 *
Mechanics	-0.093 *
Mechanical Reasoning	-0.155 **
Preferences:	
Speed of Choice	0.150 **
Interests:	
Computation	0.136 **
I.Q. Composite	-0.145 **
Socio-economic Index	-0.160 **
Social Activities:	
Dates per Week	0.121 *
Study Habits:	
Sloppy Assignments	-0.130 **
Self-report:	
Mathematics Grades	-0.109 *

*p < .05 **p < .01

NOTE: SIB Variable 296 had the following numeric code:

Took Driver Training	= 1
Did Not Take Driver Training (Not offered in their school)	= 2

With those Not Taking Driver Training having been assigned the higher numeric code value, a negative correlation between this variable and another indicates that non-takers received lower scores on the other variables.. The signs of the

Schools. Of the 861 schools, 508 offered a driver training course, 59 percent. Of the ninety-seven relationships of prime interest in this study, forty-one were significant at the .01 level and seven were significant at the .05 level. (The expected number of significant relationships by chance alone would be 4.85 at the .05 level or .97 at the .01 level if all of the ninety-seven variables were independent of each other.)

The variables that are significant at the .01 level are number offering special class for speech difficulty, number of books in the library, type of school, number of full-time teachers, number of full-time teachers who left in past year, number of full-time teachers with a bachelors degree, percent full-time certified teachers, teacher's starting salary (BS, male), numbers of tenth-, eleventh-, and twelfth-grade male students, total number of males, numbers of tenth-, eleventh-, and twelfth-grade female students, total number of females, numbers of tenth-, eleventh-, and twelfth-grade males and females, total number of males and females in all grades, offering commercial or distributive diplomas, industrial, trade or vocational diplomas, and agricultural diplomas, number of tenth-grade male dropouts, number of tenth-grade female dropouts, number of twelfth-grade male dropouts, number of twelfth-grade female dropouts, percent on half-day sessions, type of residence in school area, percentage school support from local sources, percentage of American Indians, percentage of Negro pupils, community facilities available including public library, 4-H club, Boy Scouts, etc., Girl Scouts, etc., Future Nurses Club, Future Farmers of America, Future Teachers of America, and availability of an auto mechanics course.

The variables that are significant at the .05 level are offering special class for low IQ or mentally retarded, offering special class for mathematics difficulty, concerts, average age of pupils homes, community facilities available including community theater and public swimming facility, and region of U.S. in which school is located. The forty-eight significant correlation coefficients are presented in Table 2.

Discussion

The organization and patterning of the significant relationships fit into a general theory and relate to the prior research in such a way that these relationships appear to form meaningful results.

Interpretation of the results indicates that the

Table 2.—Correlation of Selected TALENT Variables with the GSC variable Driver Education

Project TALENT Variable:	Correlation Coefficient
Speech difficulty	-.091 **
Number of books in Library	-.136 **
Type of school	+.215 **
Number of full time teachers	-.274 **
Number of full time teachers left last year	-.122 **
Number of full time teachers with B.S.	-.219 **
Number of full time certified teachers	-.138 **
Teachers starting salary, B.S., male	-.236 **
Amount of full time experience	-.090 **
Number of 10th, Male	-.221 **
Number of 11th, Male	-.201 **
Number of 12th, Male	-.210 **
Total number, Male	-.179 **
Number of 10th, Female	-.213 **
Number of 11th, Female	-.214 **
Number of 12th, Female	-.206 **
Total number, Female	-.180 **
Number of 10th, Male and Female	-.242 **
Number of 11th, Male and Female	-.226 **
Number of 12th, Male and Female	-.224 **
Total number, Male and Female (all grades)	-.193 **
Commercial or distributive	-.215 **
Industrial, trade, or vocational	-.200 **
Agriculture	-.090 **
Percent 10th dropout, Male	-.205 **
Percent 10th dropout, Female	-.246 **
Percent 12th dropout, Male	-.132 **
Percent 12th dropout, Female	-.149 **
Percent on half-day session	-.095 **
Type of residence in school area	+.214 **
Percent of school support: local sources	-.207 **
Percent American Indian	-.094 **
Percent Negro pupils	+.102 **
Public Library (facilities)	-.115 **
Auto mechanics	-.116 **
4-H Club	-.185 **
Boy Scouts (etc.)	-.212 **
Girl Scouts (etc.)	-.221 **
Future Nurses Club	-.162 **
FFA	-.156 **
FTA	-.128 **
Region	+.077 *
Low I.Q. or Mentally retarded	-.080 *
Math difficulty	-.081 *
Average age of pupils house	+.070 *
Concerts (facilities)	-.085 *
Community theater	-.071 *
Public swimming	-.080 *

Note: GSC Variable "Driver Education" had the following numeric Code: Offered Driver Education = 1; Did not Offer Driver Education = 2. With those Not Offering Driver Education having been assigned the higher numeric code value, a negative correlation between this variable and another indicates that non-offerers received lower scores on the other variables. The signs of the correlations are in a sense an artifact of the coding system. **p .01; *p .05

which offer driver training and schools which do not offer driver training cluster around socioeconomic status, a variable which is known to be related to automobile accidents. It is clear that socioeconomically related variables differentiate between the two groups of schools.

Interpretation of the results for the student data indicates that the variables which differentiate between students taking driver training and students who have no opportunity to take it because it is not offered in their school, cluster around academic knowledge, intelligence, socioeconomic status, and socioeconomic related variables. Since these are socioeconomically related variables, and thus related to automobile accidents, it becomes clearer why there is an observed relationship between having had driver education and fewer automobile accidents.

But it is also observed that this is not necessarily a cause and effect relationship. The work of McGuire and Kersh (10) further substantiates this argument.

Summary and Conclusion

McGuire and Kersh (10) indicate that variables other than the variable of having taken driver training or not taken driver training to be related to the ultimate criterion variable of highway accident frequency, accident severity, or violation frequency. Asher (1) found variables which differentiate between *students* taking driver training or not taking driver training though available. This research investigation of a random sample of 798 U. S. high school seniors indicated that differences exist between those students taking driver training and those not taking driver training because it was not offered at their schools. Interpretation of the data showed that the major factors differentiating these two groups were academic knowledge, intelligence, and socioeconomic status.

Investigation of a random sample of 861 United States high schools indicated that differences also exist between those *schools* offering driver training and those not offering driver training. Interpretation of the data showed that the major factors differentiating these two groups were socioeconomically related variables. Since it is known that socioeconomic status is related to automobile accidents, then it becomes apparent that part of the explanation of lower insurance rates and fewer automobile accidents among populations of persons who have had driver training can be explained in part by the fact that the schools and communities from which these persons came were different in socioeconomic status than from schools and communities where driver training was not offered. Hence it seems reasonable to infer that students who come from these schools probably would have had fewer automobile accidents even if they had not had driver training.

This research again does not test the more specific hypothesis that driver training makes no difference in the number of automobile accidents. (Research (2) on this hypothesis is in progress.) What this research does do is to demonstrate further that variables other than taking driver training or not taking driver training and offering driver training or not offering it differentiates between these groups. The results of this research extend the research of Conger, and others (4), and McGuire and Kersh (10) to a national sample.

FOOTNOTE

1. Part of this paper is based on a master's thesis submitted by the senior author to the Graduate School, Purdue University (11). The work reported herein was performed pursuant to a grant from the U. S. Office of Education, Department of Health, Education, and Welfare.

REFERENCES

1. Asher, J. W., "Do Driver Training Courses Produce Better Drivers? An Alternative Hypothesis," *Traffic Safety Research Review,* 12:(No. 1)2–6, 1968.
2. Asher, J. W., and Dodson, B., "Personal, Social, Educational, and Accident Variables Associated with Enrollment in Driver Training Courses," Joint Highway Research Project, Project C-36-59k, Purdue University.
3. *BIMED, Biomedical Computer Programs,* Dixon, W. J. (Ed.) Health Sciences Computing Facility, University of California, Los Angeles, 1964.
4. Conger, J. J.; Miller W. C.; Rainey, R. V., "Effects of Driver Education: The Role of Motivation, Intelligence, Social Class, and Exposure," *Traffic Safety Research Review,* 10:(No. 3)67–71, 1966.
5. Flanagan, J. C.; Dailey, J. T.; Shaycoft, M. F.; Orr, D. B.; Goldberg, I., *Studies of the American High School,* Project TALENT Office, University of Pittsburgh, Pittsburgh, Pennsylvania, 1962.
6. Flanagan, J. C.; Davis, F. B.; Dailey, J. T.; Shaycoft, M. F.; Orr, D. B.; Goldberg, I.; Neyman, C. A., Jr., *The American High School Student,* Project TALENT Office, University of Pittsburgh, Pittsburgh, Pennsylvania, 1964.
7. General School Characteristics, School Questionnaire, A National Inventory of Aptitudes and Abilities, Project TALENT Office, University of Pittsburgh, Pittsburgh, Pennsylvania, 1960.
8. Goldstein, L. G., *Research in Human Variables in Safe Motor Vehicle Operation: a Correlational Summary of Predictor Variables and Criterion Measures,* The Driver Behavior Research Project, The George Washington University, Washington, D. C., 1961. (Reprinted by U. S. Department of HEW. Public Health Service, Division of Accident Prevention, 1964.)
9. Guilford, J. P., *Fundamental Statistics in Psychology and Education,* McGraw-Hill, New York, 1956.
10. McGuire, F. L.; Kersh, R. C., "A Summary of an Experimental Evaluation of Driver Education in the United States," unpublished manuscript, University of California, 1968.
11. Shively, J. E., "Characteristics of High School Students Unable to Take Driver Training," unpublished master's thesis, Purdue University, 1968.
12. Stack, H. J. (Project writer), *History of Driver Education in the United States,* National Commission on Safety Education, National Education Association, Washington, D. C., 1966.

Research Briefs—Continued from page 171

Achievement via Conformance (Ac) vs. Achievement via Independence (Ai).

The hypothesis of the current investigation was that the Ai scale would be significantly superior to the Ac scale as a predictor of nonacademic accomplishment in college.

Method

The personnel of the Student Affairs Research Office at Iowa State University constructed a questionnaire for measuring extracurricular achievement in twelve areas: pre-professional, dramatics, literature, personal philosophy, science, social, athletic, political, general scholarship, music, art, and leadership. The questionnaire consisted of 120 "true–false" items—ten items for each area; each item described an extracurricular achievement which, in principle, could be checked against public record. The items were designed to maximize content validity and permit discrimination between varying degrees of accomplishment in the respective areas. The internal reliabilities of the resulting twelve scales were assessed (on a preliminary sample of college freshmen, $N = 107$) via the Kuder–Richardson formula 20. Reliability coefficients ranged from .36 to .77 (median = .56), which was deemed satisfactory for these 10-item scales.

During the spring quarter of 1969, questionnaires were mailed to three hundred Iowa State University freshmen, who had been randomly selected from a personal record file of freshman students. The students were also asked to complete CPI forms accompanying the questionnaires. The 164 students who completed and returned the questionnaire and CPI materials were shown, through a chi-square analysis, to be representative of the population of Iowa State University freshmen with respect to sex (final sample included 120 males and forty-four females), curriculum, and grade-point average.

Pearson product-moment correlations of Ac and Ai with the twelve scales of nonacademic achievement were computed and subjected to two-tailed tests of significance at the .01 level of confidence.

Results

The hypothesis of this investigation was not upheld, since the Ai scale was not significantly related to any scale of nonacademic achievement ($r = -.11$ to $.17$, $p > .01$). The Ac scale had significant positive correlations with general scholarship achievement ($r = .24$, $p < .01$) and leadership achievement ($r = .28$, $p < .01$), and a significant negative correlation with art achievement ($r = -.22$, $p < .01$).

At least two alternative conclusions were proposed on the basis of these results: Either the Ai scale does not sufficiently tap the motivational factors underlying extracurricular achievement or nonacademic achievement is not as "unstructured" as one would suppose.— *James D. Evans,*[1] *Iowa State University.*

FOOTNOTE

1. Now at Geneva College.

REFERENCES

1. Gough, H. G., *California Psychological Inventory,* Consulting Psychologists Press, Palo Alto, California, 1960.
2. Locke, E. A., "The Development of Criteria of Student Achievement," *Educational and Psychological Measurement,* 23:299–307. 1963.

FIELD NEWS

Willard E. North

Central Missouri State College

THE NEBRASKA CONFEDERATION in Early Childhood Education (EPDA–TTT) brings together a variety of pre- and post-doctoral scholars and staff in programs scattered across the country and concerned with the variety of ethnic conclaves, socioeconomic groups, and educational programs. The purpose of the various programs in the Confederation are: (1) to discover ways of meeting the educational needs of children 3 to 8, particularly those "locked out" by one kind of deprivation or another; (2) to meet their needs while respecting the communities and culture that also educate them; and (3) to create, in institutions of higher learning, Early Childhood teacher-training programs staffed by competent, knowledgeable persons. Eligible participants include: college teachers whose work includes training of Early Childhood teachers; pre-doctoral candidates in a variety of fields; those in leadership positions at local and state levels.

For further information write to Larry Freeman, 346 Andrews Hall, University of Nebraska, Lincoln, 68508.

* * * * * *

AN INFORMATION OFFICE established to compile and disseminate news originating from the National Network of Educational Laboratories and Research and Development Centers was opened October 1 in Denver, Colorado. The information office will serve as a base from which information about the activities under way within each center and laboratory can be accumulated and then channeled to key educational audiences. In addition, it will facilitate the flow of information of the activities within the centers and laboratories among their researchers and staff members. A newsletter will be developed to aid this effort.

E. Joseph Schneider, a staff member at the Center for the Advanced Study of Educational Administration at the University of Oregon, Eugene, will be director of the information office. The information office will be funded by participating members of the national network of laboratories and centers. Included in the latter group are a vocational and technical educational research center and a center for occupational education.

The eight university-based Research and Development Centers are working to develop prototypes for facilitating learning or for improving institutional functioning. The fifteen laboratories, less research-oriented and more concerned about development, take the prototypes produced by the centers and other agencies and then develop working models for actual use within the nation's schools.

Last year a major portion of all laboratory expenditures went for development of comprehensive instructional systems for elementary education. The development of instructional components also received a large share of the laboratory's budget. Altogether, over half of their funds went to the development of instructional and learning systems.

The federally funded centers are located at the University of Oregon; University of California, Berkeley; University of California, Los Angeles; Stanford University, Palo Alto, California; University of Texas, Austin; University of Wisconsin, Madison; The Johns Hopkins University, Baltimore, Maryland; the University of Pittsburgh, Pennsylvania.

The laboratories, partially supported by federal funds, are located in Portland, Oregon; Berkeley, California; Los Angeles, California; Albuquerque, New Mexico; Austin, Texas; Atlanta, Georgia; Durham, North Carolina; Charleston, West Virginia; Philadelphia, Pennsylvania; Newton, Massachusetts; New York City; Syracuse, New York; St. Louis, Missouri; Kansas City, Missouri; and Minneapolis, Minnesota.

The vocational and technical center is at Ohio State University, and the center for occupational education is at North Carolina State University.

* * * * * *

TWENTY-THREE PUBLIC SCHOOLS, three state departments in public instruction, and the Regional Educational Laboratory for the Carolinas and Virginia (RELCV) have formed a 3-state consortium to develop individualized curriculums for the public schools. The new organization, called the Consortium for Individualized Instructional Systems, will develop and produce curriculum materials for elementary and secondary schools.

The cooperating schools and State Departments of Public Instruction in North Carolina, South Carolina, and Virginia, will identify areas where individualized curriculum materials are needed. RELCV will develop and produce the materials in cooperation with the other Consortium members and will train teachers to use the new materials. Cooperating schools will also serve as test sites for the materials.

As its initial project, the 3-state group has contracted with RELCV to develop a mathematics curriculum, IMS–II, for grades 7 through 9. IMS–II will be an extension of the laboratory's low-cost Individualized Mathematics System (IMS) now being used by more than 8,000 pupils in grades 1 through 6 at schools in the Carolinas and Virginia.

RELCV, in cooperation with the schools and state departments, will develop and produce part of the IMS–II during the 1970–71 year. Seventh graders in the cooperating schools will pilot-test part of the Virginia high school version during the 1971–72 year. Following the testing and revision, IMS–II will be made available to other schools in the Carolinas and Virginia at cost.

The consortium will be exploring the possibilities of projects in the physical sciences, social sciences, and language arts. Dr. Robert Parlier, Superintendent of Schools in Lexington, Virginia, is chairman of the Consortium's Executive Committee.

* * * * * *

THE MAJORITY OF JUNIOR and senior high school students perceive their schools as being basically undemocratic. They report that they do not participate in making decisions which vitally affect them as individuals and citizens and that they are victims of arbitrary behavior of teachers and principals. The vast majority of students see no alternative course of action in a conflict in which they find themselves when they complain about how they are treated or when they try to do something new and different. They see most conflict resolved by the unilateral action of the teachers and administrators, and they report that such action intensifies rather than reduces tension in the school.

These are some of the principal findings of a report, "Civic Education for the 70's: An Alternative to Repression and Revolution," just released by the Center for Research and Education in American Liberties at Teachers College, Columbia University. The principal investigators are Allen F. Westin, Professor of Public Law and Government at Columbia and Center Director; John P. DeCecco, Professor of Psychology in Education at

San Francisco State College; and Dr. Arlene Richards, Associate Research Director. The study, funded by the U. S. Office of Education, aimed at formulating objectives for the reform of civic education in American schools. The investigators selected a sample of almost seven-thousand high school students in the greater New York City and Philadelphia areas. The sample included suburban and urban schools, and junior and senior high schools. In the case of one high school, all the students in attendance on the day the data were collected were interviewed. In another case, the interviewer entered a high school on a day when many students struck so that those students interviewed were a true "silent majority" who remained in classes. It is believed by the investigators that this is the first study of the "silent majority," of what current American junior and senior high school students think about the school as a sociopolitical institution. While other studies have looked at those students who led or participated in protest activities, this study looked at the school population both when the schools were quiet and when they were in turmoil.

The content analysis of the student responses yielded six major categories and issues: (1) courses and curriculum; (2) political issues outside the school; (3) infractions of legal codes outside the school; (4) school organization and governance; (5) peer conflicts outside the schools; (6) individual rights largely involving school authorities.

Issues of school governance and individual rights accounted for over 50 percent of the issues. The percentages of the other four categories were relatively small; for courses and curriculum about 13 percent; for political issues about 6 percent; for illegal acts about 10 percent, and for peer and adult conflicts outside the school about 10 percent. These findings, then, conflict seriously with the impression left by newspaper and television reports of student unrest in the high schools over racial and national political issues. Although high school students show more concern over political issues than junior high students, and suburban students more than urban students, the major source of student unrest in high schools stems from the students' lack of choice in most aspects of their daily school life and the daily obstacles they meet in the full exercise of their rights as citizens.

The remainder of the findings on how students perceive conflict and conflict resolution in the school contains major implications for school administrators. Students report that 68 percent of their conflicts are with persons in authority, only 20 percent with peers. About 46 percent perceive the conflict to be with a particular person, while 24 percent perceive the conflict with the school as an institution. Eighty percent are unwilling or unable to describe alternative choices for resolving the conflicts they face. Of those few students who perceive and describe choices, most base their choices on conviction as opposed to pure expediency.

The picture for conflict resolution is equally disappointing. Students see conflict termination (not resolved) by authority figures about 55 percent of the time; by negotiation between students and authority figures about 17 percent of the time. Many conflicts, of course, are unresolved and unending.

Nineteen percent of the incidents reported the use of force; force among peers about 6 percent; by authorities against students about 11 percent, and by students against authorities about 4 percent. According to one of the investigators, "The students see themselves to be victims of institutional violence many more times than they see themselves as instigators." Sixty-one percent see the outcomings of the conflicts as bad, only 9 percent as good. With regard to tension level, 91 percent see the tension level as not lowered, 9 percent see it lowered.

THE ESSO EDUCATION FOUNDATION has announced grants totaling $120,000 for three projects intended to improve higher education. Recipients are the American Association of Colleges, the Woodrow Wilson National Fellowship Foundation, and Stanford University.

The American Association of Colleges will use its grant to synthesize and analyze raw data available on Catholic higher education. The processed information will form the basis for the country's first data bank on Catholic higher education.

The grant to the Woodrow Wilson National Fellowship Foundation will go toward the operation of the Woodrow Wilson Administrative Internship Program. Initiated in 1967 with a grant from the ESSO Education Foundation, the Administrative Internship Program assists underdeveloped colleges, particularly black colleges in the South, by providing them with graduates from the nation's leading business and public administration schools for short term administrative assignments. This year, fifteen interns will work at thirteen small colleges as fund raisers, planners, and assistants to the presidents and business officers.

The grant to Stanford University will help underwrite the development of an interdisciplinary, international relations program for undergraduates. Departing from the traditional content of international relations programs, the new curriculum will be designed to reflect the new forces and factors governing relations among nations since World War II. Courses and seminars in the program will be taught by senior faculty from departments throughout the University, and the syllabi developed will be widely distributed to other colleges and universities.

* * * * * *

THE CATHOLIC UNIVERSITY of America has received a grant from the Ford Foundation to establish a Center for National Policy Review on Race and Urban Issues at the University Columbus School of Law. The Center will be directed by William L. Taylor, former Staff Director of the U. S. Commission on Civil Rights. It will provide non-partisan research and analysis to public interest organizations concerned with federal policy issues that affect cities and members of minority groups. Much of the Center's work will focus on deficiencies in the enforcement and implementation of federal civil rights laws and social programs, and upon the development of research for use in litigation. The Center will also provide training for law and graduate students for Catholic and other universities in ways in which public policy is formulated.

The services of the Center will be available both to national organizations in Washington, D. C., and to local and community groups throughout the nation. According to Mr. Taylor, "Many poor people and members of minorities have yet to receive tangible benefits from the numerous civil rights laws and social programs enacted during the 1960's. While those who are resisting change often have ample resources at their disposal, the people who are the intended beneficiaries of the law have a desperate need for assistance in representation. We hope that the Center will meet at least part of that need."

Among the principal concerns of the Center are the unequal distribution of resources in metropolitan areas and the barriers that central city residents encounter in their efforts to secure employment, better housing, and better public schools.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

ARMAND J. GALFO received his AB in Chemistry, EdM in Science Education, and EdD in Educational Administration from the University of Buffalo. In 1966 he was awarded a grant from the U. S. Office of Education for a year of post doctoral study and research with the Teaching Research Division of Oregon. He has taught in the public schools of New York State and Florida and joined the faculty of the College of William and Mary in 1958. In 1965–66 Professor Galfo served as Assistant Dean of the School of Education. He currently chairs the Division of Measurement and Research of the School, is the author of numerous journal articles, and is the co-author of the textbook *Interpreting Education Research.*

WILLIAM McCOLLY received his PhD from the University of California at Los Angeles. He has been a teacher of English at New Trier Township High School and on the faculty of Northwestern University and of The University of Wisconsin. He is currently on the faculty of the University of South Carolina and is a Consulting Editor of *The Journal of Educational Research.*

WILLIAM A. MEHRENS is an Associate Professor, Department of Counseling, Personnel Services, and Educational Psychology, Michigan State University. He received the BS and MEd from the University of Nebraska and the PhD from the University of Minnesota. He was previously a high school counselor, instructor at the University of Minnesota, and Assistant Director of the National Assessment of Educational Progress. BRUCE G. ROGERS is an Assistant Professor of Measurement and Statistics at the University of Maryland. He earned the BS and MA from Arizona State University and the PhD from Michigan State. His previous experience includes teaching secondary school mathematics and consulting on computer applications in educational research.

THOMAS NOFFSINGER is the Director of The Lake County Regional Program for Handicapped Children. This program deals with development of new services for handicapped children. He received the BA from Western Michigan University, and MA and PhD from George Peabody College for Teachers. VIRGINIA DOBBS is the Director of Pupil Personnel Service for the Nashville Davidson County Public Schools. She has been the Director of Psychological Services for the Nashville School System and received her MA and PhD from George Peabody College for Teachers.

JOE E. SHIVELY is a graduate student in educational research at Purdue University, specializing in statistics, measurement, and research methodology. He has done research in such areas as instructional behavior patterns, learning theory, reliability of tests, international education, and driver training. He was a recipient of a U. S. Office of Education fellowship from 1966–67 (undergraduate) and from 1967–1970 (graduate). He received his BS and MS from Purdue. WILLIAM ASHER is Professor of Education and Psychology at Purdue University and is associated with the Educational Research Center. He is a behavioral research methodologist; he was previously Research Coordinator with the U. S. Office of Education, and Professor of Educational Research at the University of Pittsburgh. He received his BA, MA, and PhD in psychology from Purdue and was a post-doctoral IBM Fellow of the Harvard Computing Center.

ALAN S. WATERMAN is presently an Assistant Professor at Rensselaer Polytechnic Institute. CAROLINE K. WATERMAN is an Assistant Professor at the State University of New York at Albany. Both received their baccalaureate degrees from the University of Rochester and their doctorates from the State University of New York at Buffalo. They have been conducting joint research in the areas of personality development during the college years and student evaluations of college experiences.

JOHN D. WILLIAMS is Associate Professor of Education in the Division of Educational Measurement and Statistics at the University of North Dakota. He received his PhD in educational measurement and research from Colorado State College. STEVEN D. HARLOW is Associate Professor of Behavioral Disabilities at the University of North Dakota. Previously he was on the staff of the Educational Psychological Clinic, University of Nebraska. He obtained his PhD in educational psychology and measurement from the University of Nebraska. DEL GAB is Assistant Professor of Education and Psychology at Dickinson State College. He received his EdD from the University of North Dakota in Teacher Education.

CHARLES H. ZASTROW is completing his doctorate in social welfare at The University of Wisconsin and is Research Director for the Dane County Social Planning Agency. His doctoral areas of specialization include evaluative research, personality theories, counseling techniques, and marriage and family systems. He received a BS in psychology and a MS in social work from Wisconsin. He is the author of "The Theory of Cognitive Dissonance" in *The Psychological Record.*

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor Curriculum and Instruction, The University of Wisconsin, Madison.

PHILIP LAMBERT, Chairman, Department of Educational Psychology and Administration, The School of Education, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1970

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Assistant Professor of Special Education, Yeshiva University, New York.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English and Education, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLASS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

CARLETON W. WASHBURNE, Director Emeritus of Teacher Education and the Division of Graduate Studies, Brooklyn College, 2248 Kent Street, Okemos, Michigan 48864.

50th Year of Publication

The Journal of Educational Researc

Dedicated to the Study of Education

VOLUME 64 JANUARY 1971 NUMBER

ARTICLES

Curriculum Theory—*James B. Macdonald* ---------------- 195
Self-Oriented Sentence Completion Responses and Reported Personal Biographical Data—*Richard G. Weigel, Fred McKinney* ---------------- 201
Comparative Achievement of Two Substandard Groups of University Freshmen—*Charles B. Keenen* ---------------- 203
Actual, Ideal, and Expected Role Concepts of Secondary Counselors—*Glenn G. Dahlem* ---------------- 205
The Effect of Ability on Concept Learning—*Joseph C. Piland, Elmer A. Lemke* ---------------- 209
An Analysis of Two Hypotheses Concerning the Relationship Between Creativity and Intelligence—*John S. Dacey, George F. Madaus* ---------------- 213
A Comparison of the Effectiveness of Discovery Versus Didactic Methods and Teacher-Guided Versus Independent Procedures in Principle Learning—*Carol Wiesner* ---------------- 217
A 5-Year Follow-Up Study of High Ability Achieving and Nonachieving College Freshmen—*Leland Smith* ---------------- 220
The Effects of a Mediational Device on Paired-Associate Learning; Learning and Retention of Sixth Grade Students—*William E. Loadman* ---------------- 223
Control of Reinforcement and Academic Achievement—*Ralph L. Shaw, Norman P. Uhl* ---------------- 226
Differentiating Motivations Underlying Vocational Choice—*Daniel Cohen* ---------------- 229
Effects of Group Counseling on Creativity—*S. Jeffrey Garfield, Helen A. Cohen, Robert M. Roth, Harris L. Berenbaum* - 235

DEPARTMENTS

Among Those Present ---------------- 240
Book Reviews—*Robert E. Clasen* ---------------- 200
Field News—*Willard E. North* ---------------- 233
Forthcoming Articles ---------------- 23
It Says in The JER ---------------- 19
New Books from the Publishers---------------- 20

(The Journal of Educational Research is indexed in Abstr.S.W., CIJE, CSPA, Curren Contents, Ed. Adm. Abst., Educ. Ind., M.H.R.Ind., PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (t May-June and July-August issues being combined). The subscription price is $10.00 a year, plus postage for all subscriptions outside the continental United States. Single copies $1.50. Second cla postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, In

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, Preside Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assista to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEW Supervisor of Editorial Services.

It Says in The JER...

CURRICULUM THEORY, says Macdonald, is presently in a state which could be called "in search of boundaries." Theorizers are seeking more satisfactory frameworks and their efforts are loosely grouped into three orientations: knowledge oriented theory; reality oriented theory; and value oriented theory (p. 195).

IMPLICATIONS FOR counseling and self concept research were discussed by Weigel and McKinney in relation to their study of college students. They used The McKinney Sentence Completion Blank and a personal biographical questionnaire to attempt to reflect characteristics of students with and without emotional problems (p. 201).

REDUCED LOAD in college may have a possible positive effect on freshmen who are admitted but have qualifications below the minimum established for entrance. Keenen found that neither quarter nor cumulative GPA's were significantly different for students receiving special guidance, sectioning, and reduced course loads as compared to students who received regular handling (p. 203).

WHAT COUNSELORS in seventy-four Wisconsin high schools believed public expectations of their activities to be did not differentiate itself from ideal role, found Dahlem. The study investigated actual, ideal, and expected role concept perceptions of the counselors (p. 205).

THIRD GRADE pupils were trained in homogeneous or heterogeneous groups on a thematic concept learning board in Piland and Lemke's study. Results indicate that ability grouping has no significant effect on concept learning under any of the effects of the variables of the experiment, and high ability Ss are better able to attain thematic concepts than medium or low ability Ss (p. 209).

AT THE HIGH IQ levels there will be a wider range of creativity than at the average and lower levels, hypothesized Dacey and Madaus. The study consisted of eighth-grade American students and secondary school Irish students. Results offered only slight support for the hypothesis (p. 213).

METHOD AND PROCEDURE FOR teaching spelling principles were tested by Wiesner. Although the results showed no significant differences for method, procedure or interaction, the discovery method and the teacher-guided group procedure showed some advantage on the delayed transfer test (p. 217).

MANY HIGH ABILITY nonachieving college freshmen might achieve if they were given early guidance and allowed the opportunity to change their academic programs, found Smith. He studied students at the University of Kentucky to determine the present status and significant differences between high ability achieving and nonachieving college freshmen who participated in a 1960–61 study (p. 220).

HIGH LEVEL mediational aids were found to significantly affect the amount and rate of original learning of sixth-grade students. Loadman found that retention data also suggest similar inferences. He attempted to determine the effect of different levels of mediational aids on a paired-associate task (p. 223).

IN THE FOUR GROUPS of children studied by Shaw and Uhl, white low socioeconomic level (SEL), black low SEL, black upper-middle SEL, and white upper-middle SEL, the locus of control scores related to success in reading in only the latter group. It was found that the higher the external score, the lower the reading score (p. 226).

CAN EXPRESSED motivations for curricular choice discriminate between college seniors in business administration and those in teacher education? Cohen found that by programming major significant factors, into a discriminant analysis, followed by a final discriminant equation, Ss can be separated into their appropriate occupational groups on the expressed motivational level (p. 229).

CREATIVITY IS POSITIVELY related to mental health, hypothesized Garfield and others. Their results indicated that improvement in group psychotherapy correlates significantly with gains in ratings by therapists on the more dynamic aspects of creativity (p. 235).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Editorial/Abstract

The State of the Art:
Curriculum Theory

Curriculum theory is presently in a state which could be called "in search of boundaries." There is no general agreement about what phenomena should or should not be included in theorizing. Further, there is disagreement among theorizers about the purpose of theorizing and at least three general groups exist: (1) those who see theory as a guide to development; (2) those who see theory as a scientific enterprise; and (3) those who see theory as an intellectual task of creating better ways to conceptualize.

What stands for theory in the minds of most practitioners today is a production model epitomized by the Tyler rationale. This technical model does not satisfy many theorizers since it avoids the essential value questions with which one must begin. Thus, theorizers are seeking more satisfactory frameworks and their efforts can be loosely grouped into three orientations: knowledge oriented theory; reality oriented theory; and value oriented theory.

Knowledge oriented work, such as the use of the concept of the structure of the disciplines and modes of inquiry, has been the greatest single development in the past decade. Essentially this approach makes epistemological questions the primary value. Much progress has been made, but many theorizers feel that epistemology is not a broad enough base for an ultimately satisfactory theory of curriculum.

Reality oriented theorizers have broadened the base of relevant phenomena and made an essentially ontological priority commitment. Some major differences arising within this group are over the nature of the basic unit of theoretical discourse, the basic purposes for theory, and the breadth of concern necessary to create adequate conceptualization.

Value oriented theorizers are essentially curriculum designers. About a half dozen designs are common today. They are variously called such names as "subject centered," "child centered," "activity," "core," "problems of living," etc. Examination of designs leads many theorizers to the conclusion that basic theoretical work must be done before adequate designs can be created.

Curriculum theorizing is one of the least understood activities in the total profession, yet it has a fundamental meaning inasmuch as it can be thought of as the essence of educational theory since it is the study of how to have a learning environment.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Curriculum Theory

JAMES B. MACDONALD
The University of Wisconsin—Milwaukee

CURRICULUM theory and theorizing may be characterized as being in a rather formative condition, for essentially there are no generally accepted and clear-cut criteria to distinguish curriculum theory and theorizing from other forms of writing in education. The present situation may be summarized by saying that curriculum theory and theorizing exists because a fair number of thoughtful and respected professional persons say they do it and that it exists. Still others refer to the work of these persons as theorizing and their efforts as theories. A reasonably knowledgeable look at the curriculum "situation" readily reveals some of the problems which create the present confusion.

To begin with, one would suspect that theory would be focused upon a clearly identified realm of phenomena. Unfortunately, this is not so in curriculum for the definitions of curriculum are as narrow as "the subject matter to be learned" and as broad as "all the experiences students have in school." Thus, writings called curriculum theory have varied on one pole from essentially epistemological statements to the other pole of statements of a "philosophy of living."

There is also some disagreement among "theorizers" about the purpose of theorizing. Among those few who give much thought to this problem there appear to be three major camps. One group (by far the largest) sees theory as a guiding framework for applied curriculum development and research and as a tool for evaluation of curriculum development.

Thus, theory becomes a springboard for prescribing and guiding practical activity in relation to curriculum. Theory in this sense functions like a philosophy in that it is not directly thought of as open to empirical validation. That this approach is not called curriculum philosophy may perhaps be so primarily because the persons who engage in it are not usually trained philosophers regardless of the fact that much of it is a form of philosophizing.

A second "camp" of ofttime younger (and far fewer) theorizers is committed to a more conventional concept of scientific theory. This group has attempted to identify and describe the variables and their relationships in curriculum. The purpose of this theory is primarily conceptual in nature, and research would be utilized for empirical validation of curriculum variables and relationships, rather than as a test of the efficiency and effectiveness of a curriculum prescription.

A third group of individuals look upon the task of theorizing as a creative intellectual task which they maintain should be neither used as a basis for prescription or as an empirically testable set of principles and relationships. The purpose of these persons is to develop and criticize conceptual schema in the hope that new ways of talking about curriculum, which may in the future be far more fruitful than present orientations, will be forthcoming. At the present time, they would maintain that a much more playful, free floating, process is called for by the state of the art.

A further interesting and sometimes complicating factor is that individuals who theorize may well operate in all three realms upon different occasions as specific professional pressures and tasks appear. Thus, any piece of curriculum theory must be looked at carefully as a specific piece of theorizing in order to assess its intent.

Huebner (9) offers an analysis of theoretical statements which is of considerable interest here. Curriculum theory he proposes can be categorized in terms of the various uses of language by theorists. Thus, he finds that there are six kinds of language used: (1) descriptive, (2) explanatory, (3) controlling, (4) legitimizing, (5) prescriptive, and (6) affiliative.

If we accept this analysis it becomes clear that "curriculum theory" varies with the intentions of theorists, as witnessed by their use of language, in any particular time and place. This may appear to be unusual in relation to the history of scientific theory at first glance, but a little reflection shows that there are similar varieties of theory in many fields. The problem that variety creates for curriculum is perhaps of a different order. It would appear that the variety is less troublesome than the confusion among theorists about the variety and of the intentions of other theorists. The result has been something like a series of theoretical exchanges which have often been at cross purposes, together with an essential lack of historical development. Instead, the historical state of the field looks much more like a set of out-of-phase cycles. It is suggested from this that curriculum theory is much in need of historical study, with the goal of untangling what Huebner referred to as the different uses of curricular language.

Kliebard (12) offers an insightful perspective on the history of curriculum issues. He highlights the ideas that curriculum has been essentially plagued by *"an ahistorical posture,"* an *"ameliorative orientation,"* and a lack of definition. He concludes that *"our basic framework and our intellectual horizons have been severely limited."* He further suggests that *"the task of the future is the development of alternative modes of thinking to the dominant 'production model' of the past 50 years."*

The production model Kliebard speaks of is that associated with Bobbit (1) and others in the early part of the 20th century and later with the Tyler (19) rationale. In Huebner's terms this is variously a *"controlling"* and *"prescriptive"* use of language. But even here the elements of control and prescription are fundamentally grounded in a technological rationale rather than in philosophical and/or scientific theory.

This technical model has been developed to its greatest sophistication by vocational education workers. An excellent recent "state of the art" review was edited by Smith and Moss (18) who summarize the process as: (1) specifying the role for which training is to be provided, (2) identifying the specific tasks that comprise the role, (3) selecting the tasks to be taught, (4) analyzing each of the tasks, (5) stating performance objectives, (6) specifying the instructional sequence, (7) identifying conditions of learning, (8) designing an instructional strategy, (9) developing instructional events, and (10) creating student and curriculum evaluative procedures and devices.

For all intents and purposes this has been what has passed for the prevailing "theory" of most curriculum workers (with variations and alterations for different areas). Many curriculum theorists, however, have not found this to be a satisfactory model for a variety of reasons; perhaps most fundamentally because the technical process begins with an acceptance of contemporary social values (thus eliminating the value question of what to teach).

The Concerns of Curriculum Theory

Philosophies of education, according to Frankema (4) are either analytical or normative. That is, they are essentially attempting to describe, discriminate, and establish meanings for terms, or they are essentially sets of statements about what should or should not be included in education and what should or should not be done during the educational process.

Curriculum theorists have found such neat categories difficult to parallel, since the concerns of curriculum at some time must be related to what is learned by persons. Thus, curriculum always has action implications with a broad directional concern for outcomes. Under these circumstances, one is always involved in assumptions and implicit (if not explicit) statements which could be classified at various times and places as ontological, axiological, and epistemological. Concern for the nature of human "being," value theory, and the nature of knowledge are intricately interwoven in action contexts. But in many ways curriculum theorizing can be conveniently categorized as oriented toward statements about knowledge, statements about the curriculum realities, and statements about valued activity.

Knowledge Oriented Statements

Undoubtedly, the most prolific group of curriculum thinkers in the past decade have been those persons concerned with knowledge. This is a reasonable and important development since the curriculum has a substance which is drawn from the accumulated cultural development of a civilization. It is indeed quite difficult to comprehend the justification of schooling outside this context. (Although whether the school is the best place to learn about it has been challenged.)

In some ways, the recent concern for knowledge has been a reactive phenomena to critics' perceptions of education as life adjustment and broad social pressures to keep up with the Russians. Whatever the motivational source of the energy in the process, a sizable number of academically talented persons have entered the curriculum arena.

Bruner (3) sounded the clarion call for the movement toward reconceptualizing the subject matter of the schools around the structure of the disciplines and the modes of disciplined inquiry. This has been picked up by persons from many disciplines and has resulted in a veritable landslide of curriculum revisions, new programs, revised and/or new materials, and in-service programs for teachers.

Essentially, as Brownell and King (2) so ably state, the rationale for the priority of disciplines lies in the assertion that man's essential nature is most reasonably fulfilled by his symbolic capacities with priority on general ideas and especially those most teachable and learnable. Thus, the curriculum needs to be fundamentally grounded in a conception of those general structures of symbolic systems which can be most communicated to and learned by others.

Schwab (17) and Phenix (16) have also been in the forefront of this general approach. Schwab has provided a fundamental analysis of the organizational, substantive, and syntactical structures of the disciplines. Phenix contributed a conceptual reorganization of the fields of knowledge with the intention of facilitating learning and use of knowledge. Vandenberg (20) edited a useful volume of readings in which he presented selected articles by a variety of persons, which help put the theory of knowledge into educational perspective.

As far as the new impetus for the disciplines and structure have moved us, there are still a sizable number of theorists who feel epistemology or knowledge is too limited a base for an adequate curriculum theory. Questions about the relevance of social, human, and personal qualities would appear to lead to broader vistas in order to cope comfortably with curriculum decisions.

Reality Oriented Statements

At the risk of misusing the concept of ontology and in a heuristic spirit, there are a number of theoretical statements which talk about the "nature of things" as they are relevant to the consideration of the curriculum. Principally, these attempts focus upon the social, cultural, and personal context and fabric which is interwoven into a complex mosaic of living and being. The curriculum thus becomes primarily a focal point for a much more fundamental concern about reality which when conceptualized can be utilized to look at curriculum.

Goodlad (5) characterized the curriculum picture in terms of two eras, the progressive era and the discipline centered era. He noted that major proponents of these eras have been heard selectively. Thus, Jerome Bruner as a major figure in the recent disciplines era has many secondary propositions which are directly correlary with concerns of the progressives; and John Dewey (as representative of progressives) warned about the tendency of some progressives to forget the disciplines. This, of course, suggests that political forces may be quite important in influencing the perceptions of scholarly work in eras of curriculum theory development.

Mann (15) made an initial foray into the relationship of politics to curriculum theory. He focused primarily upon student unrest and politics in relation to curriculum. As noteworthy as his contribution is, much more thought needs to be given and analysis needs to be carried on of the nature of the influence of changing climates of broad social and political circumstances as they impinge upon the development of curriculum theory and the selectivity of perceptions of curriculum developers when encountering this theory.

The significance of this concern should be clear. The question of whether an adequate curriculum theory can be formulated without a sophisticated awareness of political phenomena provides a dimension to theorizing that has only been noted in passing, hinted at, or broadly sketched in the past. Some theorists are beginning to wonder if these political influences may not be far more important than they generally have been thought to be.

Goodlad and Richter (6) have presented the most elaborate contemporary model for curriculum. They identify four levels of decision making: social, ideological, institutional, and instructional.

This model is predicated upon the process of rational decision-making and is an extension of the Tyler (19) rationale. However, contrary to Tyler they assume that values are beginning points not only screens to be introduced after analysis of society, learners, and subject matter, thus, avoiding one of the major criticisms of the work of Tyler (see, for example, 13).

The intent of this conceptual model is clearly to control, explain, and describe. However, limiting a conceptual model to rational decision-making processes may well rule out important descriptive and explanatory phenomena, such as that discussed by Mann, and thus weaken the long range usefulness of the model for more short term needs for control of the curriculum processes.

Macdonald (14) proposed a conceptual model which views "actions" as the central unit of curriculum theory (rather than the decisions of the Tyler-Goodlad variety). The attempt here was to explain the activity found in relevant contexts of schooling and to describe the various levels of activity that go on and at least hint at their relationships.

Johnson (10) pointed out some conceptual confusion in this model between curriculum and curriculum development. The criticism was well put and his modifications were an important clarification of the intended meaning of the conceptualization. He went on to spell out the flow of activity between curriculum development and instruction and to provide a schema for curriculum.

Again, however, it is clear that the intentions of theorists differ. Johnson clearly is seeking the kind of control that Goodlad and Richter are after whereas Macdonald's attempt was not predicated upon a control factor in terms of specific output.

Huebner (8) has been exploring at a different level. His approach has dealt with language systems and the ways in which language shapes the process of building conceptual models and/or facilitating processes and decisions in curriculum. Thus, the technical, scientific, political talk are noted in many models; but one may imply from

Huebner's discourse that the ethical and aesthetic talk about schooling has been limited, inconsistent, and of much lower priority.

It would appear then, that one central concern of theorists is identifying the fundamental unit of curriculum with which to build conceptual systems. Whether this be rational decisions, action processes, language patterns, or any other potential unit has not been agreed upon by the theorizers.

Further, it seems clear that the intentions of theorizers influence the selection of the unit. In Huebner's (8) terms, it would make sense to suggest that the intent to control predisposes selection of scientific and technical language and the conceptual system which develops reflects this initial bias. Thus, the value question has not been transcended by the curriculum models that are presently available.

Value Oriented Statements

Curriculum designs are value oriented statements. The literature is replete with suggested designs. Designs, in contrast to epistemological theories or reality oriented statements attempt to project a theoretically based pattern of experiences as desirable.

It should be noted that curriculum designs are implicit and sometimes explicit in other kinds of models. Yet the intention of designs is clearly to prescribe, legitimize, and win advocates rather than simply describe, explain, and/or control.

Over the years we have witnessed a succession of designs; from the subject centered, to broad fields, to problems of living and the child centered approach. Other prominent forms have been called the Activity Curriculum, the Core Curriculum, and the Emerging Needs Curriculum. Johnson (11), in an article on design, mentions that there appear to be from three to a half dozen current designs.

Designers, however, have not escaped the problem of the conceptual modelers. In fact, Herrick (7) suggested that in the end the designers may well be faced with the task of theorizing at the same level. There is still the problem of the basic unit around which designs are built; and the value commitment, perhaps at a different level, is central to design.

Just as rational decisions have been the predominant unit for conceptual models, designs have often utilized learning experiences as a basic unit. Other units frequently proposed have been instructional objectives, learning tasks, and functional social roles and skills.

Value priorities have generally been set in one of the basic referents of curriculum. Designers have generally opted for priority on subject matter, social phenomena, or people (learners). As arguable as this either-or position appears on a philosophical level it is extremely difficult to avoid on a practical design level since the nature of rational thought is linear and it does make a difference which one of the three one begins with. This is frequently so because the choice of priority often implies a value position about a referent that makes the definition of this referent different from what its definition would be if it came later in the set of priorities.

The problems of design are, in fact, what fostered the current interest in curriculum theory. It is reasonable to suggest that as crucial as designs are in terms of the exigencies of practical decision making, the curriculum theorist needs to do much more work before many of the design problems can be solved.

Conclusion

One may conclude that we have only touched upon the area of curriculum theory here, certainly excluding more than has been included. It is a difficult task to formalize such a diverse and wide ranging field. Yet it is an exciting venture for persons whose dispositions lead them in this direction. There is an article of faith involved which is analogous to Dewey's comment that educational philosophy was the essence of all philosophy because it was *"the study of how to have a world."* Curriculum theory in this light might be said to be the essence of educational theory because it is the study of how to have a learning environment.

REFERENCES

1. Bobbit, Franklin, *The Curriculum*, Houghton Mifflin Co., Boston, Massachusetts, 1918.
2. Brownell, John A.; King, Arthur, Jr., *The Curriculum and the Disciplines of Knowledge*, Wiley and Sons, Boston, Massachusetts, 1966.
3. Bruner, Jerome, *The Process of Education*, Harvard University Press, Cambridge, Massachusetts, 1961.
4. Frankema, William R., "A Model for Analyzing a Philosophy of Education," *High School Journal*, 2, October, 1966.
5. Goodlad, John, "Curriculum: A Janus Look," *Journal of Curriculum Studies*, 1: (no. 1) 34–46, November, 1968.
6. Goodlad, John; Richter, Maurice, "The Development of a Conceptual System for Dealing with Problems of Curriculum and Instruction," Cooperative Research Project No. 454, ED 010 064, U. S. Department of Health, Education and Welfare, Washington, D. C., 1966.
7. Herrick, Virgil E., "Curriculum Structure or Design," a mimeographed informal paper, out of print (School of Education: View of Wisconsin).
8. Huebner, Dwayne, "Curriculum Language and Classroom Meanings," *Language and Meaning*, Association for Supervision and Curriculum Development, Washington, D. C., 1966.
9. Huebner, Dwayne, "The Task of the Curriculum Theorist," mimeographed, author, Teachers College, Columbia University, New York, 1968.
10. Johnson, Mauritz, "Definition and Models in Curriculum Theory," *Educational Theory*, 17:127–140, April 1967.
11. Johnson, Mauritz, "On the Meaning of Curriculum

Design," *Curriculum Theory Network*, Ontario Institute of Education, University of Toronto, Canada, 3:3–9, Spring 1969.

12. Kliebard, Herbert, "Persistent Curriculum Issues in Perspective," *A Search for Valid Content for Curriculum Courses*, Short, Edmund (ed.), College of Education, The University of Toledo, 1970, pp. 31–41.
13. Kliebard, Herbert, "The Tyler Rationale," *School Review*, 78:259–72, February 1970.
14. Macdonald, James B., "Structures in Curriculum," Proceedings of the Conference on Curriculum Leadership, Wisconsin State Department of Public Instruction, Madison, Wisconsin, 1966 pp. 28–46.
15. Mann, John, "Politics and Curriculum Theory: An Informal Inquiry," *Curriculum Theory Network*, Ontario Institute of Education, University of Toronto, Canada, No. 5, Spring 1970.
16. Phenix, Phillip, *Realms of Meaning*, McGraw-Hill Book Co., New York, 1965.
17. Schwab, Joseph J., "Structure of the Disciplines: Meanings and Significance" in Ford and Pugno (eds.) *The Structure of Knowledge and the Curriculum*, Rand McNally, Chicago, 1964 pp. 1–30.
18. Smith, Brandon B.; Moss, Jerome, Jr., "Process and Techniques of Vocational Curriculum Development," Minnesota Research Coordinating Unit for Vocational Education, University of Minnesota, Minneapolis, April 1970.
19. Tyler, Ralph, *Basic Principles of Curriculum and Instruction*, University of Chicago Press, Chicago, Illinois, 1950.
20. Vandenberg, Donald (ed.), *Theory of Knowledge and Problems of Education*, University of Illinois Press, Urbana, 1969.

Book Reviews

ROBERT E. CLASEN
Book Review Editor

Research in Physical Education

Locke, Lawrence F. Teachers College Press, New York, May 1969, 50 pp. $1.95.

THIS BRIEF monograph is a probing critique of research in physical education. The position is taken that research in this field has had little impact on the conduct of programs of physical education and it is not likely to generate future changes. Three major shortcomings are pointed out: (1) research has become an end in itself and not a means to an end; (2) only scattered unrelated bits of information are provided by the research, much of it being contradictory and seldom dealing with the important problems that confront the teacher; and (3) the research is almost completely devoid of sound theory, thus providing little basis for educational innovations.

According to the writer, the limited theory found in the research literature is of two kinds, one involving low order generalizations which are irrelevant to curricular and pedagogical matters, and at the other extreme, highly complex theories encompassing such generic areas as "movement" or "sport." The former, according to the author, would seem to be of little use and the latter is a waste of time either for laying the foundation for further research or for the purposes of directly affecting the conduct of the program.

To fulfill its research responsibilities, the profession, according to the author, must first face the problem of improving the quality of research. The problem can perhaps be most effectively resolved by long-term sequential research training programs which are pursued jointly by training staffs and trainees. Within this framework there must be a division of effort between basic and applied research. The theories spawned by basic research need to be tested in the practical instructional situation. Research training programs of this type will of necessity utilize the research capabilities of our larger and better universities. The vast majority of teacher training institutions are not equipped to do this job. The present emphasis on research has in many instances had unfortunate effects on professional programs. Funds have been diverted to research which should have been used to upgrade already weak professional programs and the lure of research has siphoned off the best minds from teaching to the detriment of teacher training programs.

The author takes a dim view of the likelihood that basic research in physiology or related fields will produce anything relevant to the conduct of physical education. The road to relevance in physical education research will come through the concentration of the research efforts of its own people on the vital problems of the profession.

The idea that changes in the conduct of physical education programs will emerge spontaneously from teacher training colleges is, according to the author, a shortsighted point of view. There must be a concerted effort on the part of those who are knowledgeable about research to make certain that the results are made available in an understandable way to those in teacher training institutions. The author proposes the training of specialists who will serve as middlemen between research and teaching. If such a procedure is not adopted, some new organizational scheme will have to be set up to bridge this gap.

The age old questions are posed, "What questions are worth asking?" and "What kind of information is worth knowing?" The researcher has traditionally considered that he is the one to make this decision and any departure from this is considered by him to be an encroachment. The author believes this should be a joint responsibility between the researcher and the consumer. He does not speculate on the effect such a practice is likely to have on the researcher. One might ask, whether it is better to study the basic phenomena which control behavior or attempt to find answers to the pressing questions of the hour. While this is a matter upon which opinions differ, it has been the view of scientists generally that the greatest benefits to mankind have in the long run come from basic research and from the freedom of the researcher to seek answers to questions of concern to him.

In most respects the author's criticisms are justified and the recommendations sound. The monograph will be received enthusiastically by the practitioner, less so by the researcher. Nevertheless, it should be read by both.

G. Lawrence Rarick
Professor of Physical Education
University of California, Berkeley

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Self-Oriented Sentence Completion Responses and Reported Personal Biographical Data[1]

RICHARD G. WEIGEL[2] and FRED McKINNEY, University of Missouri

ABSTRACT

The relationship of responses on a self-oriented sentence completion blank and personal adjustment, as inferred from reported personal biographical data, was examined. The McKinney Sentence Completion Blank (MSCB) and a personal biographical data questionnaire were administered to 160 college students. A Self Actualizing response pattern on the MSCB was found to reflect reported personal biographical data characteristic of unselected, emotionally "healthy" students, while a Defensive response pattern was positively related to biographical data characteristic of students with emotional problems. The Immobilized Affect response pattern did not appear to reflect immobilization in the traditional sense. Implications for counseling and for self concept research were discussed.

SENTENCE completion blanks have traditionally been of value in counseling practice. Through such semi-projective instruments, the client may reveal characteristic experience and behavior patterns in terms highly individual to himself. Based on such information, the counselor often can better formulate hypotheses about the client, and subsequently use them as a basis for further exploration in counseling sessions. Such instruments, however, have typically not been designed to yield statements about the self. A particularly important class of client responses, which is related very directly to personal adjustment, has hence been neglected.

The present paper reports the use of an experimental self-oriented sentence completion blank, with a focus on its ability to yield responses related to student adjustment. In particular, it presents the relationship of responses on a sentence completion blank and reported personal biographical data. Biographical data has been shown previously to distinguish reliably between students with emotional problems and healthy students (3), and also to contribute to a prediction of academic achievement in the university (1).

Method

Instruments. The McKinney Sentence Completion Blank (MSCB) has been described in detail by McKinney (2), Schneider (4), Spillman (5), and Vajgert (6). The instrument is designed to generate statements about the self through the use of forty-six stems, all but five of which contain the first person pronoun. These stems elicit responses to frustrations, conflict, and personal identity, and allow the individual to explore indirectly and existentially various aspects of self-involvement.

The content of an individual's responses may be used directly to generate hypotheses for counseling relating to the self. In addition, the counselor may choose to score the MSCB in order to classify the client's responses into psychologically significant categories. Responses are categorized as Self Actualizing (A), Defensive (D), or Immobilized Affective (I). Responses not meeting the criteria for these categories are Non-classifiable (N). The specific classification criteria for scoring are as follows:

Self Actualizing *(A): Responses which seem to lead to need satisfaction, and show the inclination of the individual to engage in trial and error behavior or alternative behavior. The criteria include responses which show an openness to new experience, a sense of responsibility for one's actions, democratic social interests, and adaptive behavior broadly conceived.*

Defensive *(D): Responses which are likely to lead to failure or digression in need satisfaction. Specific criteria include behavior which is socially undesirable or self-defeating, maladaptive, or personally unrewarding; classic defense mechanisms, inappropriate situational responses, and evasive responses.*

Immobilized Affective *(I): Responses concerned with the collapse of adaptive capacity with attendant unresolved anxiety. Despair at problem resolution, psychosomatic reactions, and responses describing panic, fear, worry, shame, or embarrassment.*

Non-classifiable *(N): Responses which do not yield enough information to be placed in any of the above categories, or responses which could be placed on more than one of the scales.*

The number of responses in each of these categories is determined for the individual, and in addition, the ratio of A to all scorable responses is calculated A/A + D + I. McKinney (2) has reported intra-subject reliabilities of .89 after 3 days, and .58 and .51 after 14 weeks on this ratio. He also reported intra-judge reliabilities on scoring the ratio of .92 and .95, and inter-judge reliability of .88.

Reported personal biographical data were elicited by use of a 25-item true-false questionnaire. The items were derived from those found by McKinney (3) to distinguish reliably between a group of college students with emotional problems, and an unselected (and it was inferred "healthy") control group.[3] The score on the questionnaire was calculated by determining the number of items answered in the same direction as the "healthy" group. This phenomenological instrument was chosen, rather than a more objective measure of personal data, on the assumption that it is not reality, but perceived reality which most affects individual adjustment.

The study was designed to determine the relationship of responses on the self-oriented sentence completion blank (MSCB) and personal adjustment, as inferred from responses of personal biographical data from the questionnaire.

Procedure and Results. The MSCB and the questionnaire were completed by 160 undergraduate students at the University of Missouri. The Ss included 115 men and forty-five women, of whom ninety-two were freshmen, forty-seven sophomores, twelve juniors, and nine seniors. Median age was 19. A number of groupings of Ss were made, based on MSCB category scores, MSCB ratio scores, and questionnaire scores. Three groups were selected on the basis of MSCB category scores:

A Those fifteen Ss scoring highest on the A scale, and hence whose responses best reflected Self Actualizing actions or attitudes.

D Those fifteen Ss scoring highest on the D scale, and hence whose responses best reflected Defensive self-defeating actions or attitudes.

I Those fifteen Ss scoring highest on the I scale, and hence whose responses best reflected Immobilized Affect.

Using an analysis of variance and orthogonal comparisons, both the A and I groups were found to have significantly higher scores on the questionnaire than the D group ($F\text{'s} = 4.43$, $p < .05$). Groups A and I did not differ significantly ($F = 0$).

Two groups were selected on the basis of the $A/A + D + I$ ratio scores of the MSCB. These were the Ss who had the fifteen highest and fifteen lowest ratio scores. Using a t-test, a trend was noted for the high-ratio group to have higher scores on the questionnaire than the low-ratio group ($t = 1.91$, $.10 > p > .05$).

Two groups were selected on the basis of scores on the 25-item questionnaire. The high-score group was composed of those twenty Ss scoring most like McKinney's (3) emotionally "healthy" group, whereas the low-score group was composed of those twenty Ss scoring most like his group with emotional problems. The high-score group was found to respond to the MSCB with significantly higher ratio scores ($t = .93$, $p < .05$), significantly more A responses ($t = 2.48$, $p < .05$), and significantly fewer D responses ($t = 3.46$, $p < .01$) than the low-score group, using t-tests. No differences were observed between groups in the number of I responses ($t = .32$).

Finally, Fisher Product Moment Correlations were determined between the A, D, I, and ratio scores of the MSCB, and scores on the questionnaire. Questionnaire scores were found to be significantly correlated +.23 with A responses, −.26 with D responses, and +.24 with ratio scores. No significant correlations were produced with I responses.

Discussion

The results indicate that responses to a self-oriented sentence completion blank are related to personal adjustment, as inferred from reported personal biographical data. Positive relationships were noted between a A orientation (A responses and high-ratio scores) on the McKinney Sentence Completion Blank and reported personal biographical data similar to an emotionally "healthy" group of students. Conversely, positive relationships were noted between a D orientation (D responses) and reported personal data similar to a group with emotional problems. It should be noted, however, that individuals responding to the MSCB in a manner reflecting affect which was seen as incapacitating (i.e., many I responses) were not found to be more similar in reported personal biographical data to a group with emotional problems than to a "healthy" one. Presumably, some amount of healthy affect at times may accompany A, and not be immobilizing. A conceptual tightening of the I category of the MSCB to exclude non-immobilizing affect would appear to be called for.

Overall, the MSCB appears to have promise as an instrument for counseling and research purposes. It may be particularly useful in assessing specific variables of client orientation toward the self which have relevance to personal adjustment, such as futurity, degree of self-identity, level of maturity, and realism.

FOOTNOTES

1. Based in part on a paper presented at the annual meeting of the Oregon and Washington Psychological Associations, 1965.
2. Now at Colorado State University.
3. Instrument available from senior author.

REFERENCES

1. Harrington, C., "Forecasting College Performance from Biographical Data," *Journal of College Student Personnel,* 10:156–160, 1969.
2. McKinney, F., "The Sentence Completion Blank in Assessing Student Self-actualization, *Personnel and Guidance Journal,* 45:709–713, 1967.
3. McKinney, F., "Case History Norms of Unselected Students and Students with Emotional Problems," *Journal of Consulting Psychology,* 11:258–269, 1947.
4. Schneider, R. N., "Some Relationships Between 'Self-management' and Self-Experiencing," unpublished PhD dissertation, University of Missouri, Columbia, 1962.
5. Spillman, A., Jr., "'Self-management' as Related to Certain Aspects of Self-experiencing in a Female Population," unpublished MA thesis, University of Missouri, Columbia, 1964.
6. Vajgert, D. G., "Internal Consistency of the McKinney Sentence Completion Blank," unpublished MA thesis, University of Missouri, Columbia, 1964.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Comparative Achievement of Two Substandard Groups of University Freshmen

CHARLES B. KEENEN
Germanna Community College, Fredericksburg, Virginia

ABSTRACT

Two groups of freshmen were admitted to the University of Florida whose qualifications were below the established minimums. The students in one group were routinely handled. The second group received special guidance, sectioning, and reduced course loads. The two groups were matched on test scores, and their achievement, as measured by grade point average (GPA), was compared over two quarters. Neither the quarters nor cumulative GPA's yielded a significant difference between the two groups. A possible positive effect of the reduced load is suggested by the study and considered in the discussion.

FOR A number of years the University of Florida met the increasing pressures on freshman enrollment by a gradual strengthening of admissions requirements. Following the entrance of the freshman class in the fall of 1963, it was determined that class size would be controlled by a straightforward limit on the number of qualified students who would be permitted to enter, rather than any additional increase of requirements.

The basic minimum academic requirements since the fall of 1963 have been (1) a "C" average in academic subjects in high school and (2) a score in the top 40 percent of high school seniors in the state on the Florida State-wide Twelfth Grade Testing program tests. In the summer of 1967 these requirements were modified permitting exceptions based on other evidence indicating the probability of success for no more than 5 percent of the entering freshman class.

Students admitted under this provision have been below the minimum on only one of the two basic academic requirements; therefore, they should not be conceived of as seriously deficient academically nor representative of any disadvantaged or culturally deprived group. These students have taken the regular freshman courses and have not been in a special preadmission program such as reported on by Grant and Angleman (1). As these students were unquestionably marginal, they did afford the opportunity to investigate several devices hoped to improve or enhance their achievement and success.

Sample

In September 1967 a group of forty students was admitted as exceptions in that they failed to meet one of the basic admission requirements. Fourteen students had test scores below the minimum, and twenty-six had high school averages less than "C." Under the exception provision in June 1968, forty-four students were admitted. Eighteen presented substandard test scores, and twenty-six were deficient on high school average. There were twenty-nine men and eleven women in the September group, while the June entrants were thirty-five men and nine women. The N's in the analyses vary from these totals as one of the students in the September group withdrew during the first quarter (injury) but returned in the second, and six of the September entrants did not return for the second quarter. There was a loss, also, of four students from the first to second quarter in the June group.

Method

Both groups of students admitted as exceptions to the minimum academic requirements were matched on the Florida Twelfth Grade (FTG) Testing program's total scores. Assuming that the variation of independent influences on freshman achievement is essentially the same between several quarters as within one quarter, then the fact that these groups entered in different quarters is not felt to prejudice this study.

The students who entered in September were given no special treatment; they were not identified in any way, no grouping was done, and no extra counseling or advisement was given. For all practical purposes, they were as any other entering freshmen and have been permitted to continue in this fashion. The students who entered in June were assigned to the same freshman English (3 hours) and Logic (2 hours) sections, and this arrangement was continued under the same instructors for two quarters. During the first quarter there was an extra period a week in each of these courses, and students found deficient in reading skills were required to participate in a program at the Reading Clinic. The two selected instructors closely followed the total progress of each student, encouraging and implementing added counseling when deemed appropriate. The students in the June group were permitted to, and did, carry a reduced load during the first quarter; however, they carried the minimum (12 hours) or more during their second quarter.

The two groups were to be compared on their achievement as measured by their University GPA on completion of the first quarter's work, on the second quarter's work, and on the cumulative record of the two quarters. The special conditions afforded the June group were not extended beyond the first two quarters of attendance.

Results

The correlation between FTG test scores and University GPA for this study sample was found to be .16. This low magnitude is not unexpected in so far as the study groups were composed of students deficient on only one measure or requirement of test score or high school average. The quarter and cumulative averages were obtained from the official records of the University of Florida Registrar's Office, and t-tests were computed. The results are presented in Table 1.

Table 1.—September 1967 and June 1968 Students Compared on Mean University GPA's

	September, 1967			June, 1968			
	N	MEAN	S.D.	N	MEAN	S.D.	t
1st Quarter GPA	39	1.597	.744	44	1.859	.610	1.69
2nd Quarter GPA	34	1.579	.721	40	1.630	.526	.34
Cumulative GPA	33	1.664[*]	.565	40	1.755	.476	.82

[*] The higher cumulative GPA was caused by the loss of the six students from the first quarter and the general trend of lower averages in the second quarter.

These results indicate that there was no significant difference in the achievement of the September and June groups in any of the three periods investigated, as t-tests failed to attain the required levels. Therefore, the special guidance, reduced load, and assignment to the same sections did not enhance the achievement of the June group as measured by University GPA. In only one period, the first quarter of attendance, did the difference in GPA even approach significance suggesting a possible causative factor which will be explored in the discussion.

Summary and Discussion

In the fall of 1967 the University of Florida changed from a trimester to a quarter calendar. Existing University data, correlations, and formulas were based on the trimester system and students' performance under it, and, therefore, are of questionable validity for application to the present study under the quarter calendar. As such information was available, some subjective consideration seemed worthwhile.

Predicted University GPA's (for a trimester calendar) are computed in a routine manner in the admissions process for each entering freshman. As is usual in academic prediction equations, previous performance, in this case the high school average, carries the greatest weight. While the two groups in the present study were matched on test scores, the group entering in September, which received no special attention, had a higher mean high school average of 2.097 compared to 1.956 for the June group. When this high school average was weighted in predicting University average, the September group yielded a mean predicted University GPA of 1.6745 and the June group one of 1.5647 (a difference significant beyond the .01 level, $t = 3.41$). In terms of actual achievement, the September group's mean University GPA was not only lower than predicted, but also lower than the June group's. The June students outperformed their own prediction and both the prediction and achievement of the September group. This suggests that the special efforts invested in the June group may have had a greater influence than has been able to be determined statistically at this point.

The factor mentioned in reporting the results that seemed to be exerting a strong influence on achievement was the number of hours carried. During the first quarter of attendance in which there was the largest mean University GPA difference, the September group carried a mean of 13.4 hours, while the June group, as one of their special conditions, carried a mean of 9.3 hours, a difference of 4.1 hours. In the second quarter, the mean hours carried were 13.8 and 12.8 for September and June groups respectively, and the mean University GPA difference was .05. As the other benefits afforded the June group (except for hours carried and the extra class session) continued through the second quarter, the strong influence of hours carried is implied.

Two groups of freshmen with matched test scores, whose qualifications failed to meet part of the minimum established academic requirements, were admitted to the University of Florida, one group in September of 1967 and the other in June of 1968. The September group was given no special attention. All of the June group were assigned to the same sections in two required freshman courses for two quarters and received an extra class per week in these courses during the first quarter, carried a reduced load during the first quarter, were specially advised by the instructors in the two courses, and participated in the Reading Clinic program when deemed necessary. A comparison of University GPA for two quarters' work, individual and cumulative, failed to yield any significant difference in level of achievement. In the discussion, information on predicted and actual achievement was considered which suggested more positive implications of the study. It would appear that additional studies of this type involving added variables and controls are warranted.

REFERENCE

1. Grant, C. W.; Angleman, S. W., "A Pre-admission Program for Students with Low Academic Promise," *Journal of College Student Personnel,* 9:320–325, 1968.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Actual, Ideal, and Expected Role Concepts of Secondary Counselors

GLENN G. DAHLEM
Pattonville (Missouri) Public Schools

ABSTRACT

This study investigated actual, ideal, and expected role concept perceptions of seventy-four Wisconsin secondary school counselors. The 50-item checklist of Schmidt (5) was used as the research instrument. Expected role, i.e., what counselors believed public expectations of their activities to be, did not differentiate itself from ideal role. Significant (.001) differences were found on ten of the fifty work activities contained in the checklist. Counselors indicated a desire to participate in more professional duties than they were then doing. Addition of more research activities was the paramount role alteration wanted.

SEVERAL researchers, Dunlop (2), Lund (3), Roth (4), Schmidt (5), Tennyson (9), Wasson and Strowig (10), have investigated actual and ideal role concepts held by secondary school counselors. They have reported discrepancies between actual and ideal roles held and have speculated regarding the causes of these. No study, however, has investigated disparities among counselors' own perceptions of actual, ideal, and expected professional roles.

Dunlop (2) studied attitudes toward the counseling function among groups of California counselors, administrators, teachers, parents, high school seniors, and counselor-educators in 1963. He administered a five degree of response questionnaire to Ss who rated possible counselor duties in seven responsibility areas.

Although he found all groups agreeing on some counselor duties, Dunlop uncovered wide disagreements from group to group on many of the duties. For instance, parents and students expected the counselor to be an advice-giver, while the other groups rejected this notion. Apparently, various publics view the counselor in light of their own frame of reference.

Lund (3) compared the perceptions of the ideal counselor role held by counselors, principals, and teachers in all Minnesota high schools having full-time counselors in 1961–62. A Q-sort technique of twenty-eight duties, to be judged "appropriate" or "inappropriate" by Ss, was used. Lund found that all three groups had similar role perceptions, with counselors and principals in closer agreement than either was with teachers. He found, as did Schmidt (5), that presence of so-called "inappropriate" counselor duties in actual work situations was apparently the result of factors other than role perceptions. He found no relationship between school size and differences in counselor role perceptions among any of the groups. While Dunlop also found principals closer in agreement with counselors than teachers were, Lund did not attribute his counselor-teacher differences to an orientational difference, as Dunlop did.

Roth (4) compared actual and ideal roles of 267 counselors and sixty-two teacher-counselors in Wisconsin senior high schools to a role model based on substantial agreement of twenty-four Wisconsin counselor-educators. He used an instrument similar to the one employed in this study. Roth concluded that actual roles of counselors and teacher-counselors were unique to given employment situations. He found that a desirable ideal role model exists, in terms of major functions. He recommended that a role model be flexibly structured because of differing school situations and counselor abilities.

Schmidt (5) studied actual and ideal counselor role concepts among forty counselor-principal pairs in Missouri high schools. He used a 50-item Q-sort of possible counselor duties as his instrument. He found counselors and principals much closer in agreement regarding counselor role than is popularly believed, and discovered that counselors' differences in role concepts were about as great as those of principals. He attributed 60 percent of the disparity between actual and ideal role to factors other than differing role concepts of the two groups.

Tennyson (9) investigated the differential between percent of work time devoted by Missouri secondary school counselors to various functions, as contrasted to percent of time they felt should be so devoted. The percent of time devoted to several duties in each of four assistance categories, namely, students, teachers, administrators, and research, was studied. Significant differences between actual and ideal role appeared in the first two categories.

Tennyson found counselors more in agreement on how their time should be spent than on how it was being spent. The greatest amount of time was spent in the area of helping students, with "counseling" receiving about half of that. The counselors would have spent more time on all student activities except individual appraisal and occupational-educational information, on which they would spend less. Time spent on assistance to teachers was less than that advocated by most authorities in the field, Tennyson maintained. The counselors felt more time should be spent helping teachers acquire guidance skills and adapting class instruction to pupil needs, and less in discussing "problem" children.

Wasson and Strowig (10) investigated forty Wisconsin high school counselors, comparing twenty counselors (isolates) from 1-counselor schools to twenty from multi-counselor schools. They sought to determine by means of interviews whether significantly different role concepts existed among isolated as opposed to non-isolated counselors. The investigators found that isolates had greater job satisfaction, and felt their faculties were more supportive than did non-isolates. The non-isolates, however, identified more with their professional leadership and did more professional reading.

The present study investigated similarities and discrepancies among actual, ideal, and expected roles as counselors themselves report them. Here, counselors indicated their perceptions of their professional role as follows:

actual—what counselors state they are doing in their work,

ideal—what counselors state they would do in their schools if they could structure an ideal work situation for themselves,

expected—what counselors state they believe the residents of their school districts think counselors should do in their schools.

Reported actual and ideal roles were compared. In addition, these roles were compared with the reported expected role, to determine if this third role concept is really different from the other two in the counselors' own perceptions.

Shertzer and Stone (7) defined counselor role in Parsonian terms, as a set of complementary expectations that result in behavior. They discussed role concepts of various publics (teachers, pupils, etc.) of the counselors from a frame of reference of expectations of each public. The present study approached expected role in a similar manner, but from the standpoint of counselor perception of public opinion, rather than direct study of community beliefs. This approach, while it does not objectively measure counselor role expectations among the general public, does assess what counselors believe these to be. Thus, it becomes possible to contrast counselors' expected role perceptions with their reported actual and ideal role concepts. This, in turn, may make possible certain insights into the effect, if any, that expected role perceptions have upon performance of various counselor duties, and provide directions for future research.

Method

The population studied was the seventy-four public secondary school counselors of a 3-county area of southern Wisconsin. They represented 100 percent of all junior and/or senior high school counselors employed in the area during the 1967–68 school year.

Table 1.—Prefacing Instructions for Checklist Administrations

Actual Role	Ideal Role	Expected Role
"The following is a list of possible duties performed by a high school counselor. Indicate, by checking the appropriate column, which you do or do not do in your own job situation. Then, if you wish, circle those duties which are absolutely necessary and essential to your work."	"If you had the privilege of doing only the things you wanted to do in your job situation, which of the following would you do? Indicate, by checking the appropriate column, which things you would and would not do. Then, if you wish, circle those duties which are absolutely necessary and essential to your work."	"What things, to the best of your knowledge, do you feel that the people of your school district think you should do in your job as counselor? Indicate, by checking the appropriate column, which things the people think you should and should not do. Then, if you wish, circle those duties the people consider absolutely necessary and essential to your work."

The research instrument was the 50-item checklist of Schmidt (5). This instrument lists fifty possible counselor duties in each of five service areas. According to Schmidt, twenty-five of the items state "proper" counselor duties, and twenty-five state "improper" duties. The items were subjected to an elaborate Q-sort process by him, to eliminate ambiguities as much as possible.[1]

Counselors received three administrations of the instrument, each fitted with a different set of prefacing instructions. These administrations sought to measure actual, ideal, and expected role concepts. The order of administration was mathematically varied throughout the group, to neutralize any effects of set caused by recollections of previous responses. On their checklists, counselors indicated whether they do or do not (actual role), would or would not (ideal role), and should or should not (expected role) perform each of the fifty duties. Table 1 lists prefacing instructions for the three checklist administrations.

Respondent data were analyzed in fifty 3 x 2 chi-square matrices, one for each checklist item (6). If the assumptions of chi-square could not be met (expected frequencies below six) in a given matrix, it was broken up into three 2 x 2's, so that Fisher's exact test could be substituted (8). In order to compare one role directly to another, the method of exact partitioning of chi-square contingency tables was used (1). Only items found significant at the .001 level were considered significant.

Results and Discussion

Ten items were found significant in the comparison of actual to ideal role, and eight of these items were significant in comparison of actual to expected role. No items were significant in comparison of ideal to expected role. The significant items were (see footnote 1): 02, 03, 07, 19, 27, 28, 29, 33, 37, and 39. Of these, 07 and 37 were significant for actual to ideal role only. Table 2 lists respondent totals for the significant items.

Expected role did not emerge as a separate and distinct concept in this study. No significant differences between ideal and expected role were found. What counselors ostensibly would do in an ideal work situation corresponded closely with what they indicated the public expects they should do. These findings do not necessarily refute existence of counselor expected role, rather, they establish that for this particular investigative approach, it did not differentiate itself.

In considering the significant differences found between actual and ideal role, it is noted that for all ten significant items, ideal response totals were more highly affirmative than actual totals. It may be concluded, then, that counselors would perform more professional duties in their professed ideal work situations than they indicated they were actually performing.

Table—2. Respondent Totals for Significant Checklist Items

Signif.	Actual Role		Ideal Role		Expected Role	
Item #	do	not	would	not	should	not
02	41	25	60	5	56	8
03	5	59	21	44	13	48
07	22	43	45	20	34	26
19	34	31	54	11	52	11
27	47	16	62	4	62	2
28	35	30	60	5	53	11
29	22	42	42	21	31	30
33	33	30	54	11	45	18
37	24	39	43	19	37	26
39	31	32	55	10	49	13

Looking at the ten significant items, five of the ten concern various research activities (items 02, 03, 07, 19, 33), two involve working with youth-concerned outside groups (29, 39), two concern enhanced scope of services offered to special categories of students (28, 37), and one involves working with fellow staff members (27). Apparently, the greatest single alteration these counselors would make in their professional roles is expansion of research activities. Because counselors consistently included more professional activities in ideal roles than were admitted to be in actual roles, a favoring of diversification of duties is identifiable among the population studied.

FOOTNOTE

[1] The items are as follows:

00. Counsel non-school adults
01. Supervise audio-visual programs
02. Conduct follow-up studies of graduates
03. Conduct research on teaching methods
04. Tabulate data of research studies for the school
05. Record absences or tardies
06. Confer with probation or parole officer
07. Plan community occupational survey
08. Develop local norms for standardized tests
09. Discuss common student problems with the principal
10. Work with individual teachers to develop their counseling skills
11. Maintain occupational and educational information file
12. Counsel students with personal or social problems
13. Identify exceptional students
14. Prepare transcripts on students entering college
15. Administer an individual test to all students
16. Do research on school system's special education program
17. Direct research studies of the curriculum
18. Interpret guidance program through speeches
19. Direct research studies to evaluate guidance services

20. Section or group students for teachers
21. Conduct case conferences
22. Carry on work of principal in his absence
23. Routinely interview truants and/or their parents
24. Interpret test results to individual students
25. Regularly provide teachers personal information on students
26. Make suggestions to teachers on teaching methods
27. Discuss purposes of guidance program in staff meetings
28. Conduct exit interviews with all drop-out students
29. Consult with professional staffs of Y.W.C.A., Boy Scouts, etc.
30. Secure glasses, shoes, etc., for needy students
31. Act as liaison between school and police on student investigations
32. Directly use state employment services with all students
33. Use research to identify common student problems
34. Make periodic report of your work to the principal
35. Consider with a teacher a problem he has with a student in class
36. Interview students referred by principal
37. Organize special activities for new students
38. Routinely counsel parents of failing students
39. Serve on civic committees and study groups on youth
40. Plan schedule of classes for each student
41. Maintain cumulative records
42. Interview students referred by teachers
43. Prepare news articles on outstanding students
44. Record notes of interviews with students and parents
45. Collect autobiographical material on each student
46. Make item analyses of teacher-made classroom tests
47. Interview teachers whom students feel have problems
48. Counsel parents concerning problems of a student
49. Counsel students on vocational and educational planning

REFERENCES

1. Bresnahan, Jean L.; Shapiro, M. N., "Exact Partitioning of Chi Square Contingency Tables," *Psychological Bulletin*, 6:252–262, 1966.
2. Dunlop, Richard S., "Educators, Parents, and Students Assess Counselor's Role," *Personnel and Guidance Journal*, 44:1024–1028, 1965.
3. Lund, Duane R., *A Study of Counselor, Principal, and Teacher Perceptions of the Secondary School Counselor*, unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1962.
4. Roth, Richard R., *A Role Analysis of Wisconsin Secondary School Counselors*, unpublished doctoral dissertation, The University of Wisconsin, Madison, 1968.
5. Schmidt, Lyle D., "Concepts of the Role of Secondary School Counselor," *Personnel and Guidance Journal*, 41:600–605, 1962.
6. Senders, Virginia L., *Measurement and Statistics*, Oxford University Press, New York, 1958.
7. Shertzer, Bruce; Stone, Shelley, "The School Counselor and His Publics: A Problem in Role Definition," *Personnel and Guidance Journal*, 41:687–695, 1963.
8. Siegel, Sidney, *Non-Parametric Statistics*, McGraw-Hill, New York, 1956.
9. Tennyson, W. Wesley, "Time, The Counselor's Dilemma," *Personnel and Guidance Journal*, 27:127–133, 1958.
10. Wasson, Robert; Strowig, R. Wray, "Professional Isolation and Counselor Role," *Personnel and Guidance Journal*, 45:451–460, 1965.

New Books from the Publishers

Disadvantaged

Hilsheimer, George von. *How to Live with your Special Child: A Handbook for Behavior Change* (Washington, D. C.: Acropolis Books, 1970), 272 pp. Cloth $7.50.

Horn, Thomas D. (ed). *Reading for the Disadvantaged: Problems of Linguistically Different Learners* (New York: Harcourt, Brace, and World, Inc., 1970), xiii + 267 pp. Paper $4.50.

Johnson, Marjorie Seddon, and Roy A. Kress (eds.). *Reading Difficulties: Classroom and Clinic* (Philadelphia: Temple University, 1970), 132 pp. Paper $4.

Love, Harold D. *Parental Attitudes toward Exceptional Children* (Springfield, IL: Charles C. Thomas, Publisher, 1970), ix + 167 pp. Cloth.

General

Edwards, A. D. *The Changing Sixth Form in the Twentieth Century* (New York: Humanities Press, Inc., 1970), xii + 116 pp. Cloth $3. Paper $1.50.

Field, Dick. *Change in Art Education* (New York: Humanities Press, Inc., 1970), ix + 138 pp. Cloth $3.75.

Fischer, Joseph. *The Social Sciences and the Comparative Study of Educational Systems* (Scranton, PA: International Textbook Company, 1970), xiv + 533 pp. Cloth $10.

Friedrichs, Robert W. *A Sociology of Sociology* (New York: The Free Press, 1970), xxiii + 429 pp. Cloth $11.95.

Hass, Glen, Kimball Wiles, and Joseph Bondi. *Readings in Curriculum*, second edition (Boston: Allyn and Bacon, Inc., 1970), xvi + 708 pp. Paper $6.50.

Hu, C. T. (ed.). *Aspects of Chinese Education* (New York: Teachers College Press, 1970), 95 pp. Paper $3.25.

Keeney, Barnaby C. (Chairman). *National Endowment for the Humanities*, Fourth Annual Report (Washington, D. C.: U. S. Government Printing Office, 1970), iv + 77 pp. Paper $.45.

Munro, Margaret. *The Psychology and Education of the Young* (New York: American Elsevier Publishing Company, Inc., 1970), x + 294 pp. Cloth $4.95.

Mussen, Paul H. and Mark R. Rosenzweig (eds.). *Annual Review of Psychology* (Palo Alto, CA: Annual Reviews, Inc., 1970), xi + 674 pp. Cloth $10.

Stensaasen, Svein. *Interstudent Attraction and Social Perception in the School Class* (Boston: Universitetsforlaget, 1970), 282 pp. Paper $9.

Travers, Robert M. W. *Man's Information System* (Scranton, PA: International Textbook Company, 1970), vi + 175 pp. Paper $3.25. Cloth $6.25.

Wolff, Max and Alan Rinzler. *The Educational Park: A Guide to its Implementation* (New York: Center for Urban Education, 1970), 56 pp. Paper $.60.

Research and Measurement

Cronbach, Lee J. *Essentials of Psychological Testing, third* edition (New York: Harper and Row, Publishers, 1970), xxx + 752 pp. Cloth $10.50.

Holt, Robert T. and John E. Turner (eds.). *The Methodology of Comparative Research* (New York: The Free Press, 1970), ix + 419 pp. Cloth $8.95.

Schoer, Lowell A. *Test Construction: A Programmed Guide* (Boston: Allyn and Bacon, Inc., 1970), x + 176 pp. Paper $2.95.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

The Effect of Ability Grouping on Concept Learning[1]

JOSEPH C. PILAND[2] and ELMER A. LEMKE, Illinois State University

ABSTRACT

An experiment was conducted to investigate effects upon training and transfer of: ability grouping, intelligence, sex, and temporal test. Ninety-six third-grade pupils were trained in homogeneous or heterogeneous groups on a thematic concept learning board. Subjects were subsequently tested individually under an immediate or delay transfer condition. Results indicate that (a) ability grouping has no significant effect on concept learning under any of the effects of the variables of the experiment, and (b) high ability Ss are better able to attain thematic concepts than medium ability or low ability Ss. The nonsignificant effect of ability grouping is seen as a significant finding in light of its present emphasis in the school setting. An effect found significant in the controlled setting may have an effect in the school, but we doubt that an effect that is not significant in the controlled setting will suddenly work in the school.

ONE OF the most popular of numerous plans proposed by educators for organizing instructional groups is that of ability grouping. Reviews of research on ability grouping (1, 2, 3, 4, 5, 6, 7, 8) indicate that the research has been characterized by diversity of experimental method and statistical design yielding highly paradoxical results. Reviewers have been especially critical of the internal validity of the experiments.

The purpose of this experiment was to determine whether grouping, studied under laboratory conditions in which internal validity was rigorously controlled, would provide a condition for more efficient concept learning. It was felt that if grouping is a significant variable for school practice, surely it would be revealed under these conditions.

Method

Subjects. The sample consisted of sixteen each of high intelligence males, high intelligence females, medium intelligence males, medium intelligence females, low intelligence males, and low intelligence females selected from the pupils enrolled in the third grade of the twelve Bloomington, Illinois, public schools. Subjects were classified by intelligence scores attained on the Kuhlman Anderson Test, Booklet CD, as either high (109 or above), medium (96–103), or low (90 or below). They were randomly assigned to homogeneous training and heterogeneous training conditions, within each school, based upon these IQ scores.

Materials. The stimulus materials forming the display consisted of photographs, each of which represented four attributes with one of three defining characteristics for each attribute. The four attributes and their three defining characteristics were:

1. Superman: standing-flying-disguised.
2. Friends of Superman: Mr. Olsen-Mr. White-Miss Lane.
3. Scene: office-dock-cave
4. Number of Bombs: one-two-three

The concept attainment display consisted of a 5 feet by 5 feet ordered arrangement of eighty-one cards. The concepts imbedded in the display were randomly selected from all possible two-attribute conjunctive-type concepts.

Procedures. Upon entering the room, Ss were seated in chairs approximately 3–5 feet from the concept learning board. Ss were asked whether they could see all the pictures on the board. Next Ss were told they were going to play a game and were cautioned to pay close attention to the instructions. After reading the instructions, Trainer 1 (T_1) provided Ss the opportunity to attain two-dimensional conjunctive concepts.

Following the training session T_1 remained with two Ss in the room where the training had taken place while Trainer 2 (T_2) went to another identically equipped testing room. T's then stationed the S to be tested after an 18-minute delay, the delay subject (Sd), outside the testing room. Sd was permitted to read material of his choosing. Each T then tested the S assigned to immediate transfer, the immediate subject (Si), for 18 minutes to determine the number of concepts that could be attained. Upon completion of testing, Si was dismissed and Sd was similarly tested. The immediate and delay temporal condition was included only to facilitate the experiment—without it four T's would have been

Table 1.—Means of Main Effects of Grouping (Homogeneous–Heterogeneous), Intelligence (High–Medium–Low) and Sex (Male–Female) in the Training Task

Mean	Homogeneous Mean	Heterogeneous Mean		Mean
High = 3.84	4.75	3.62	Male	
	3.25	3.75	Female	
Medium = 1.87	3.50	1.62	Male	Male = 3.62
	1.25	1.25	Female	Female = 2.96
Low = 0.87	0.26	0.37	Male	
	1.25	1.25	Female	
Homogeneous	3.65	Heterogeneous 2.92		

required on the transfer tasks. A test for the main effect of trainers was not significant, and the trainer effect is not included in the transfer analysis.

Results

The results of the training concepts were derived from a 2 × 3 × 2 factorial design with main effects of grouping (homogeneous—heterogeneous) × intelligence (high—medium—low) × sex (male—female). Table 1 presents the means for the training task and Table 2 presents the ANOVA Summary.

The results of the transfer concepts were derived from a 2 × 2 × 3 × 2 grouping (homogeneous—heterogeneous) × sex (male—female) × intelligence (high—medium—low) × temporal test (immediate—delay), factorial design. Table 3 presents the means for the transfer design, and Table 4 presents the ANOVA Summary.

In the analysis of both training and transfer tasks intelligence was the only factor found to be statistically significant. The Duncan multiple range test revealed significant differences ($P<.01$) in concept learning performance between Ss of high intelligence and Ss of low intelligence and between Ss of high intelligence and Ss of medium intelligence. There was no significant difference between Ss of medium intelligence and low intelligence although the mean performance scores were higher for medium intelligence Ss than for low intelligence Ss.

Discussion

Studies of ability grouping have yielded highly paradoxical results. Reviews of the studies gen-

Table 2.—ANOVA Summary of Grouping (Homogeneous–Heterogeneous), Intelligence (High–Medium–Low) and Sex (Male–Female) in the Training Task

Source	Ss	df	MS	F
Between Cells	189.86	11	17.26	
A (Group)	5.51	1	5.51	.56
B (Intelligence)	146.02	2	73.01	7.46 *
C (Sex)	4.59	1	4.59	.47
AB	3.39	2	1.69	.17
AC	8.76	1	8.76	.89
BC	18.81	2	9.40	.96
ABC	2.77	2	1.38	.14
Within Cells	821.37	84	9.77	
Total	1011.24	95		

*$p < .01$

Table 3.—Means of the Main Effects of Grouping (Homogeneous–Heterogeneous), Temporal Test (Immediate–Delay), Intelligence (High–Medium–Low), Sex (Male–Female) in the Transfer Task

	Homogeneous 6.79		Heterogeneous 5.66			
	Immediate Mean	Delay Mean	Immediate Mean	Delay Mean		Mean
High = 11.46	12.25	15.25	12.00	13.00	Male	
	7.00	11.75	9.25	11.25	Female	
Medium = 4.53	7.25	6.50	3.75	3.25	Male	Male = 6.85
	5.25	4.00	3.50	2.75	Female	Female = 5.60
Low = 2.68	2.75	4.50	0.50	1.25	Male	
	1.25	3.75	3.00	4.5	Female	
	Immediate = 5.64		Delay = 6.81			

Table 4.—ANOVA Summary of Grouping, Temporal Test, Intelligence and Sex for the Transfer Task

Source	Ss	df	MS	F
Between Cells	1683.46	23	923.36	
A (Grouping)	30.37	1	30.37	.83
B (Temporal Test)	32.66	1	32.66	.89
C (Intelligence)	1372.14	2	686.07	18.84*
C (Sex)	37.5	1	37.5	1.03
AB	6.	1	6.	.16
AC	21.93	2	10.96	.30
AD	42.67	1	42.67	1.17
BC	51.51	2	25.75	.70
BD	2.04	1	2.04	.05
CD	70.18	2	35.02	.96
ABC	7.64	2	3.82	.10
ABD	0.04	1	0.04	.00
ACD	5.39	2	2.69	.07
BCD	3.14	2	1.57	.04
ABCD	.18	2	.09	.00
Within Cells	2621.5	72	36.41	
Total	4304.95	95		

*$p < .01$

erally question the internal validity of the experiments—the confounding of variables with the grouping condition. It was felt by the authors that conceptual learning tasks in an experimental setting would reveal the significance of the grouping condition and bring some clarity to an area fraught with controversy.

The importance of the results of the experiment lies in the fact that no statistical significance could be found among the main or interaction effects with the exception of the stratifying variable (intelligence). The failure to find a significant difference in learning efficiency for either homogeneous or heterogeneous grouping conditions is consistent with well-controlled research which also indicates that neither homogeneous nor heterogeneous groupings improve learning efficiency. The significant effect of the stratifying variable (intelligence) is consistent with previous research.

FOOTNOTES

1. The authors wish to express their gratitune for cooperation received from the staff and school systems of Bloomington, Illinois, and the laboratory school of Illinois State University. The data on which this

paper is based were included in the major author's dissertation in educational administration.

2. Now President at Lincoln Trail Junior College.

REFERENCES

1. Borg, W. R., "Ability Grouping in the Public Schools," *Journal of Experimental Education*, 34: 1–97, 1965.
2. Cornell, E. L., "Effects of Ability Grouping Determinable from Published Studies, The Grouping of Pupils," *Thirty-fifth Yearbook, Part I, National Society for the Study of Education*, Public School Publishing Company, Bloomington, Illinois, 1936.
3. Ekstrom, R. B., "Experimental Studies of Homogeneous Grouping: A Critical Review," *School Review*, 69: 216–226, 1961.
4. Keliher, A. V., "A Critical Study of Homogeneous Grouping," *Contributions to Education No. 452*, Bureau of Publications, Teachers College, Columbia University, New York, 1931.
5. Miller, W. S.; Otto, H. J., "Analysis of Experimental Studies in Homogeneous Grouping," *Journal of Educational Research*, 21: 95–102, 1930.
6. Passow, H. A., "The Maze of the Research on Ability Grouping," *The Educational Forum*, 26: 281–288, 1962.
7. Shores, J. H., "What Does Research Say About Ability Grouping by Classes?" *Illinois Education*, 53: 169–172, 1964.
8. Wyndham, H. S., "Ability Grouping: Recent Developments in Methods of Class-Grouping in the Elementary Schools of the United States," *Educational Research Series, No. 31*, Melbourne University Press, Melbourne, 1934.

New ---- **ON TO THE CLASSROOM**

A Guide for Preschool Teachers in Head Start-Type Programs of Compensatory Education

Edited by ROBERT E. CLASEN

200 pages $7.25 Hardcover and $5.75 Softcover

ON TO THE CLASSROOM deals with typical problems common to teachers of disadvantaged preschool children and contains unique suggestions for understanding and meeting the needs of these youngsters. The chapters are based on papers by well-qualified professors and professionals from the preschool education field which were originally presented to a group of Head Start teachers needing help in the various areas covered. The editor says, "Since these works were extremely useful to one group of teachers, they should be useful to others."

The book begins with a chapter which defines "culturally deprived" and offers a frame of reference for the thoughts and ideas presented in the remainder of the book. Each chapter was selected by Dr. Clasen on one criterion:

Does it contain information which our experience has shown that teachers need?

The chapters speak for themselves:

Creating a Learning Environment (numerous hints are given on how this learning climate can be created)

The Teacher, The Child and Head Start (the needs of children and a teacher's awareness are discussed)

Speech Language Acquisition and **Language and Head Start** (deal with language diagnosis and teaching strategies)

From a Teacher's Point of View (a humorous and heart-rending day to day account of organizing, canvassing, and parent programming in Heart Start, plus the happenings in a Head Start classroom from the first class day to the last—all taken from a teacher's log with her commentary and suggestions)

A Conversation with a Head Start (A.D.C.) Mother (reveals what the mother of a Head Start child experiences)

Programming for Parents (offers surprising views on what this is all about)

ORDER FROM—Dembar Educational Research Services, Inc.
Box 1148 • Dept. W • Madison, Wis. 53701

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

An Analysis of Two Hypotheses Concerning the Relationship Between Creativity and Intelligence[1]

JOHN S. DACEY and GEORGE F. MADAUS, Boston College

ABSTRACT

The purpose of this study was to test McNemar's hypothesis that ". . . at the high IQ levels there will be a very wide range of creativity, whereas as we go down to the average IQ and down to the lower levels the scatter for creativity will be less and less." The sample consisted of 867 American eighth-grade students from twenty-three junior high schools in three states plus 182 Irish secondary school students from suburban Dublin. In the American sample, divergent thinking was measured by four tests adopted from those of Torrance. The Lorge–Thorndike was used as a measure of intelligence. Selected tests from the Minnesota Tests of Creative Thinking and the AH 4 General Test of Intelligence were administered to the Irish sample. In both samples the IQ distributions were divided into three equal ranges. A one-way analysis of variance and a correlational analysis were performed. Results offered only slight support for the McNemar hypothesis. The authors suggest three possible explanations of these findings. Their major hypothesis has to do with a number of validity problems associated with creativity tests.

REACTIONS to Getzels and Jackson's investigation of the relationships between creativity and intelligence (6, 7) sparked theoretical speculation as to the nature of this relationship at several IQ levels. McNemar, in his 1964 presidential address to the American Psychological Association, was one of the first to hypothesize that the relationship would differ at various ranges. He argued that ". . . at the high IQ levels there will be a very wide range of creativity, whereas as we go down to average IQ, and on down to lower levels, the scatter for creativity will be less and less. Having a high IQ is not a guarantee of being creative; having a low IQ means creativity is impossible" (13:879). Guilford describes the scatterplot for divergent production and IQ as triangular, that is:

> *It is such that those with high IQ may be found almost anywhere along the range on a DP (divergent production) test; those who are low on the DP test can also be almost anywhere on IQ, but those high on the DP test have a high probability of being above average on IQ (10:166).*

In other words, both theorists assert that the variance between the two constructs is greater at the high level of IQ than at the low and intermediate levels. This study investigates this assertion, referred to hereafter as the McNemar hypothesis.

Barron further defines the relationship by suggesting that creativity is positively correlated with IQ at the lower but not the higher levels of IQ.

> *In brief, for certain creative activities a specifiable minimum IQ probably is necessary in order to engage in the activity at all. But beyond that minimum . . . creativity is not a function of intelligence as measured by IQ tests (2:22).*

This study also examines Barron's hypothesis.

Review of the Literature

Although not designed to test either hypothesis, Ripple and May's (17) data offer relevant evidence. They identified groups of seventh graders at three levels of IQ: low (72–90), middle (96–110), and high (116–133), with thirty students in each group. It should be noted that these ranges were not equal. The correlations between three of their nine (Guilford-type) creative-thinking tests and IQ (the Otis Quick-Scoring) were considerably lower at the low range of IQ than at the high range, thus contradicting Barron's hypothesis. On the other hand, four of the creative-thinking tests correlated higher at the low range of IQ than at the high range, thus supporting the hypothesis. There were no differences in relationship between the low and high range of IQ for the two remaining tests.

In a study which was designed to test the McNemar hypothesis, Ginsberg and Whittemore (9) used Mednick's (15) Remote Associates Test (RAT) as a measure of creativity and the ACT (American College Testing Program) as a measure of scholastic aptitude. The sample was comprised of 292 University of Nevada students. After a visual check of their scatterplot revealed no curvilinearity, they computed the RAT deviations about the RAT-on-ACT, split the abscissa at six locations (N's ranged from 43 to 63), and compared variances between columns. They concluded that their data did not support McNemar's hypothesis. Since they did find a slight trend of an increasing RAT deviation from lower to higher ACT scores, they replicated their design, using RAT and Australian Council for Educational Research (ACER) intelligence test scores of 107 Australian 15- and 16-year-old boys. On the basis of the two studies, they concluded that "the hypothesis is refuted" (9:139).

Mednick and Andrews (14) also used the RAT and several measures of IQ to examine the McNemar hypothesis (although the present authors contend that their correlational approach is really only applicable to the Barron hypothesis). They found that the correlations between the RAT and IQ generally declined as the average IQ level increased (from seventh graders to college freshmen), to some extent a confirmation of Barron's hypothesis. However, Mednick and Andrews acknowledged that this study did not overcome the problem of restricted range. They therefore administered the RAT and Scholastic Aptitude Test (SAT) to 1,211 college freshmen. The data were arranged in five equivalent SAT bands in order to cancel the effect of restricted range. From correlations which for the most part were not significant, they concluded that:

> *Among these college students, there is no differential decline in the relationship as the level of intelligence increases; if anything there was an opposite tendency. Therefore, the suggestion that creativity and intelligence are relatively independent processes among the very bright and more closely related at lower levels of intelligence is not supported in this particular instance (14:430).*

The choice of the RAT by Ginsberg and Whittemore (9) and Mednick and Andrews (14) makes their findings suspect, if Guilford (10) is correct in arguing that the RAT is more likely a measure of convergent production than of divergent production.

In summary, most studies have not found support for either the McNemar or Barron hypotheses. However, these studies were less than adequate tests of the hypotheses, in that they studied populations homogeneous in IQ, e.g., college students, used small samples, and/or examined unequal IQ ranges. The present study avoids each of these shortcomings, and follows the suggestion of Mednick and Andrews that "It would be worthwhile to examine this relationship in data from populations with a wider range of intelligence: e.g., a typical high school" (14:431).

Procedure

The hypotheses were tested on three separate samples, designated "Eastern," "Midwestern," and "Irish." Subjects in the Eastern sample were 867 eighth-grade students enrolled in twenty-three junior high schools in New York, Pennsylvania, and New Jersey. In each of the twenty-three schools, two eighth-grade classes were randomly selected for testing. The average chronological age was approximately 13 years, and there was an almost exactly equal number of boys and girls in the sample.[2]

The IQ test used in this study was the Lorge-Thorndike Intelligence Test, Level 4, Form A, Verbal (11). The creative battery was made up of four subtests: Imagination, Asking Questions, Guessing Causes, and Guessing Consequences. In order to compute a total creativity score, scores on each of these subtests were standardized to a mean of 100 and a standard deviation of 20. The four standardized scores were added together and the total divided by four. A description of each of the subtests can be found in Ripple and Dacey (16).

Scores on the IQ test were ranked from high to low; the range was 79–148. In order to examine equivalent ranges, three IQ bands of 20 points were selected: 85–104, 105–124, and 125–144. There were two Ss below this range, and four above it. These six Ss were eliminated from the study.

The Midwestern sample consisted of 583 high school freshmen and sophomores from a suburb south of Chicago. The school serves lower-middle and middle-class families with a preponderance of blue collar workers. Creativity scores were derived in the same way as in the Irish sample. For the measure of intelligence, the total converted score of the School and College Ability Test (SCAT), administered upon admission to the high school, was available on all Ss. Ranges on the SCAT were: low, 261–277; middle, 278–294; and high, 295–311. Twenty-five Ss' scores fell outside and were eliminated.

The Irish sample was composed of 182 students from suburban Dublin. The students were at the same level as sophomores in American high schools, and had a median age of 14.8. Total creativity scores were derived by summing scores from four subtests of the Minnesota Tests of Creative Thinking, scored according to Yamamoto's (20) procedures. The intelligence test used was the AH 4 Group Test of Intelligence (1). Ranges on the AH 4 were: low, 34–53; middle, 54–73; and high, 74–93. Three Ss' scores fell outside this range and were eliminated.

Results

To test the McNemar hypothesis that the variance between intelligence and creativity scores is greater at the high than at the low and middle levels of intelligence, Bartlett's test for homogeneity of variance was used. Table 1 presents the variance within each intelligence range, the range size, the number of Ss in each range, and the X^2 associated with Bartlett's test for each sample. For the Eastern sample, the χ^2 was 9.80, which is significant at the .05 level. However, the variance is highest at the middle level of intelligence, which contradicts the hypothesis. The X^2 for the Midwestern sample was 1.03, which is not significant at the .05 level and 1.62 for the Irish sample, which is not significant at the .05 level. Therefore, McNemar's hypothesis is not supported by any of these data.

Table 1.—Variances, Ranges, N's and the χ^2 Associated with Bartlett's Test for Homogeneity of Variance for Each of the Three Samples

	EASTERN	MIDWESTERN	IRISH
X^2 FROM BARTLETT'S TEST	9.80*	1.03	1.62
LOW IQ VARIANCE	532.25	3,951.38	2,743.02
IQ RANGE	85-104	261-277	34-53
N	191	160	27
MIDDLE IQ VARIANCE	672.28	3,478.64	2,923.57
IQ RANGE	105-124	278-294	54-73
N	545	306	117
HIGH IQ VARIANCE	653.83	3,865.11	4,052.59
IQ RANGE	125-144	295-311	74-93
N	131	118	35

* Significant at the .05 level.

To test Barron's hypothesis that the positive relationship between intelligence and creativity is higher at the lower levels of intelligence than at the higher levels, correlations between the two were computed at three equal ranges of intelligence. Table 2 presents the correlations between the IQ test and the creativity measures for the Eastern, Midwestern, and Irish samples.

Table 2 shows that for the Eastern sample, the correlations between IQ and creativity are higher at low and middle ranges of intelligence than at the high range; while this difference is significant at the .05 level, the amount of variance explained by the correlations is so minuscule as to offer only slight support to Barron's hypothesis. For both the Irish and Midwestern samples, none of the correlations of IQ with creativity were significantly different from zero. Therefore, these data offer no support to the Barron hypothesis.

Discussion

Despite the incorporation of factors designed to improve on the methods of previous studies, the results of this study offer virtually no support for either the McNemar or Barron hypotheses. The variance was not greater at the high intelligence level than at the low level. Furthermore, nearly all the correlations are nonsignificant at the low as well as the high levels of IQ in all three samples. The data in the Eastern sample tended only slightly in the direction hypothesized by Barron. Although the results of this study are similar to those of the studies reviewed, the interpretation offered by the present authors of these findings differs substantially from those of the other authors. The following alternative explanations are suggested.

First, the construct of creativity may be so unstable at the adolescent stage that nonsignificant correlations with intelligence are to be expected. Dacey and Ripple (4) studied the test-retest reliabilites for a 5-month period, using the same measures as those in our Eastern sample and eighth-grade students as Ss. The correlations were positive and low to moderate. Yamamoto (20) found low to moderate 2-month test-retest correlations, using early adolescents. He found considerably higher correlations for college seniors. Dacey and Ripple (4) also correlated scores on creativity measures with nine measures of personality and found low correlations in most cases. The same variables have been found to correlate much higher in adult and some preadolescent populations. If creative ability is unreliable during the adolescent years, then both hypotheses may be inappropriate for this age group.

Second, the McNemar and Barron hypotheses may be incorrect at any age level, i.e., the variances may be as great at the low levels of intelligence as at the high levels. This would be so if there were as many highly creative persons with low intelligence as there are with high intelligence.

There is substantial evidence that many more creative acts are performed by persons with high IQ's than by persons with moderate or low IQ's (e.g., 2, 10, 18). Despite the fact that intelligence appears to be a necessary condition, it is probably not a sufficient condition, since many persons of high IQ give little evidence of creativity in their work. Donovan's distinction between the "intellectual" and the merely "intelligent" individual is relevant:

> *Minds of penetrating intelligence provide the society with its experts and professionals in many fields, but these minds tend to be employed in living* off *rather than* for *ideas. They do not have the free-wheeling, critical, creative, and speculative bent of mind that marks the intellectual* (5: 39).

Third, it may be that the currently popular measures of divergent thinking (including those

Table 2.—Correlations at Each Level of IQ for the Three Samples

	EASTERN	MIDWESTERN	IRISH
TOTAL IQ CORRELATIONS	.26*	.06	.08
LOW IQ CORRELATIONS	.16*	-.03	.12
MIDDLE IQ CORRELATIONS	.17*	.04	-.03
HIGH IQ CORRELATIONS	.03	.07	-.14

* Significant at the .05 level.

used in this study) are not a fair test of either hypothesis, because they are of dubious validity. Among the causes of this doubt are questions concerning the reliability of the measures (3), relevancy of scoring and other methods factors (8, 12), testing environment (19), and the absence of any studies of predictive validity and almost any of concurrent validity. Whether or not the McNemar and Barron hypotheses can ever be tested with paper-and-pencil divergent thinking tests depends upon the solution of these problems, and more probably upon the development of more creative methods of measuring the construct.

FOOTNOTES

1. The research reported in this paper was supported by a grant from the School of Education, Boston College, Chestnut Hill, Massachusetts.
2. The authors wish to thank Dr. Richard Ripple of Cornell University's Learning Structures Project for his cooperation in the use of these data.

REFERENCES

1. *AH4 Group Test of Intelligence*, National Foundation for Educational Research in England and Wales, London, undated.
2. Barron, F., "The Dreams of Art and Poetry," *Psychology Today*, 2: (no. 7) 184, 1968.
3. Dacey, J. S., "Programmed Instruction in Creativity and Its Effects on Eighth Grade Students," unpublished doctoral dissertation, Cornell University, Ithaca, New York, 1967.
4. Dacey, J. S.; Ripple, R. E., "Relationships of Some Adolescent Characteristics and Verbal Creativity," *Psychology in the Schools*, 6: (no. 3) 321–324, 1969.
5. Donovan, J. D., "Creating Anti-intellectuals?" *Commonweal*, 81: (no. 2) 37–39, 1964.
6. Getzels, J. W.; Jackson, P. W., "The Creative Adolescent: A Summary of Some Research Findings," *The Third (1959) University of Utah Research Conference on the Identification of Creative Scientific Talent*, Taylor, C. W. (ed.), University of Utah Press, Salt Lake City, 1959, pp. 46–57.
7. Getzels, J. W.; Jackson, P. W., *Creativity and Intelligence: Explorations with Gifted Students*, John Wiley and Sons, Inc., New York, 1962.
8. Getzels, J. W.; Madaus, G. F., "Creativity," in Ebel, R. J. (ed.), *Encyclopedia of Educational Research (4th ed.)*, Macmillan, New York, in press.
9. Ginsburg, G. P.; Whittemore, R. G., "Creativity and Verbal Ability: A Direct Examination of Their Relationship," *British Journal of Educational Psychology*, 38: (no. 2) 133–139, 1968.
10. Guilford, J. P., "*The Nature of Human Intelligence*" McGraw-Hill, New York, 1967.
11. Lorge, I.; Thorndike, R. L., *The Lorge-Thorndike Intelligence Test, Level 4, Form a, Verbal Battery* Houghton-Mifflin, New York, 1954.
12. Madaus, G. F., "A Cross-cultural Comparison of the Factor Structure of Selected Tests of Divergent Thinking," *The Journal of Social Psychology*, 73: 13–21, 1967.
13. McNemar, Q., "Lost: Our Intelligence. Why?" *American Psychologist*, 19:871–882, 1964.
14. Mednick, Martha T.; Andrews, F. M., "Creative Thinking and Level of Intelligence," *Journal of Creative Behavior*, 1: (no. 4) 428–431, 1967.
15. Mednick, S. A., "The Associative Basis of Creative Process," *Psychological Review*, 69:220–232, 1962.
16. Ripple, R. E.; Dacey, J. S., "The Facilitation of Problem Solving and Verbal Creativity by Exposure to Programmed Instruction," *Psychology in the Schools*, 4: (no. 3) 240–245, 1967.
17. Ripple, R. E.; May, F. B., "Caution in Comparing Creativity and IQ," *Psychological Reports*, 10: 229–230, 1962.
18. Roe, A., "Personal Problems and Science," in Taylor, C. W. (ed.), *The Third (1959) University of Utah Research Conference on the Identification of Creative* Creativity and IQ," *Psychological Reports*, 10:229– *Scientific Talent*, The University of Utah Press, Salt Lake City, 1959, pp. 202–212.
19. Wallach, M.; Kogan N., *Modes of Thinking in Young Children*, Holt, Rinehart and Winston, New York, 1965.
20. Yamamoto, K., *Experimental Scoring Manual for the Minnesota Tests of Creative Thinking and Writings*, Kent State University, Bureau of Educational Research, Kent, Ohio, 1964.

SCRIBBLE WRITING

• Have you tried "scribble writing" in your kindergarten or first grade? The results are often amazing. On any type of paper without lines have the children write a story in scribble lines, which may be straight, looped, zigzag, or up and down. Since the child is not stymied by choosing only words he knows how to spell and then laboriously writing them out, there is a great freedom and speed in scribble writing. Afterwards have each child read his story to the class. It works every time. The child reads his story fluently, turning the paper to continue the reading of his composition. Even the reluctant, shy ones will participate.—From *Catholic Elementary Education News.*

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

A Comparison of the Effectiveness of Discovery Versus Didactic Methods and Teacher-Guided Versus Independent Procedures in Principle Learning

CAROL WIESNER, University of Connecticut

ABSTRACT

A 6-week experimental program was written to teach six basic spelling principles to sixteen sixth-grade classes. The classes were randomly assigned to four experimental treatment groups—each of which used a different combination of method and procedure. One immediate and one delayed posttest were given to determine spelling achievement in terms of retention, transfer, and problem-solving abilities. Analysis of variance procedures used to analyze the results showed no significant differences for method, procedure, or interaction. Several of the results approached significance, however, as the discovery method and the teacher-guided group procedure showed some advantage on the delayed transfer test. Implications for further research are suggested.

ONE OF the most controversial issues in the field of learning today is concerned with the value of the discovery approach. Under the leadership of Jerome Bruner (2), the champions of the cause have attributed many benefits to learning by discovery—such as a shift from extrinsic to intrinsic rewards, a way of conserving memory, and a technique for learning the heuristics of discovery.

An opposing position has been taken by David Ausubel (1) who has tried to break through what he has termed the "elaborate mystique" surrounding the discovery method. A number of other noted researchers, among them Cronbach (3) and Wittrock (6), voiced the same kinds of concerns at a conference on discovery learning which was sponsored by the U.S. Office of Education and held at Stanford University in 1965.

The general feeling which seemed to emerge from this conference was one of optimism in regard to there being a place for discovery within the school curriculum. Accompanying it, however, was a feeling of dismay due to the lack of research needed to define the proper place and function for the discovery approach. A number of suggestions were made at the conference in regard to the type of research needed to fill the void.

Problem and Hypotheses

The purpose of this study was to make use of the guidelines given at the conference in an attempt to determine the value of the discovery method in the area of principle learning. Its effectiveness was to be evaluated in terms of retention, transfer, and problem solving—both immediate and delayed. The comparison method was the didactic method which the discovery method had been expected to replace.

The study was also designed to test the effectiveness of two different classroom procedures used in conjunction with both the discovery and didactic methods. A teacher-guided group procedure was selected because of its widespread use; it was to allow for maximum group discussion. It was compared with an independent procedure which has been incorporated in many of the new curriculum programs which feature the discovery approach. This study, then, was designed to test the following null hypotheses:

1. *There is no significant difference in the learning of general principles between students taught by the discovery method and those taught by the didactic method.*
2. *There is no significant difference in the learning of general principles between students taught by a teacher-guided group procedure and those taught by an independent procedure.*
3. *There is no significant interaction of discovery and didactic methods with teacher-guided group or independent procedures.*

Related Research

A number of studies have been carried out on the discovery method in both the laboratory and classroom during the past 40 years. Although the results are quite contradictory, several general trends seem to emerge. Didactic treatment groups seem to learn faster and perform better on practiced tasks, as evidenced by their superior scores on immediate retention tests. Although discovery treatment groups take more time for initial learning, they seem to retain more material over a longer period of time. They seem to do consistently better on all immediate and delayed transfer tests. Those studies which have investigated problem-solving performance also show an advantage for discovery groups.

Methodology

Learning Task. The principles which were selected were basic spelling principles involved in adding endings to base words. They were selected because they are fundamental to every spelling program and would therefore be appropriate for determining the implications of the discovery method for a particular academic discipline. This curriculum area was also selected because it has received little attention in regard to incorporating the discovery approach even though educators such as Horn (5) and Hildreth (4) have suggested the use of the discovery method as one step toward improving spelling programs.

Treatments. The combination of two methods with two procedures resulted in the use of four experimental treatments. Groups A and B used the discovery method with Group A using a teacher-guided group procedure and Group B using an independent procedure. Groups C and D used the didactic method with Group C using a teacher-guided group procedure and Group D using an independent procedure.

Materials and Procedures. An experimental program was written for the study. It consisted of a series of six spelling lessons and was designed to run for 6 weeks. A new spelling principle was introduced in each lesson. The experimental treatments differed only in the manner in which the principle was taught—with each group using the combination of method and procedure to which it had been assigned via random procedures. The same materials and techniques were used for all groups for the practice, testing, and review parts of the weekly lesson.

Testing Program. The testing program used in the study included one standardized test and two special tests. The standardized test was the spelling subtest of the Iowa Test of Basic Skills which had been given as part of the regular testing program in the schools. Analysis of its scores indicated no significant differences between the treatment groups in regard to spelling achievement before the experimental program began.

The two special tests which were developed were used as posttests. One was given at the conclusion of the experimental program, the other 6 weeks later. Each of the posttests consisted of three parts: Part A was designed to measure retention and consisted of words which had been practiced in the lessons; Part B was designed to measure transfer and consisted of words which had not been practiced but to which the principles applied; Part C was designed to measure problem-solving and consisted of three sets of sample words for which the student had to derive the principle involved in order to apply it to new words. Content validity for the two tests was established by a panel of seven judges. Reliability was checked by giving the test to four classes which were not otherwise involved in the study. The resulting coefficients of equivalence were significant beyond the .01 level.

Population and Sample. The study was conducted in the public school system of Glastonbury, Connecticut. All of the sixteen sixth-grade classes in the system were involved in the study with four classes randomly assigned to each treatment. Students were not told that they were taking part in an experiment and teachers were cautioned against the use of any procedures which might suggest the experimental nature of the program or bias the results. Analysis was carried out on the scores of eighty-seven Ss in each treatment group for a total of 348 Ss.

Statistical Treatment. The major statistical technique used to analyze the data was multiple classification analysis of variance. The assumptions of homogeneity of variance, random sampling, and a normally distributed population had been met in the study. A 2x2 model was used to study the effects of the two independent variables and their interaction. The generalized formula for this model is represented by the following equation: $(Y_{mpi} - \bar{u}) = A_m - B_p - C_{mp} - E_{mpi}$. It states that a person's deviation (Y_{mpi}) from the population mean ($\bar{u}$) is considered to be a function of the two factors of method (A_m) and produre (B_p), the interaction of method × procedure (C_{mp}), and the error term (E_{mpi}). The error term consisted of the within mean square for the two main effects and their interaction because the type of design used constituted a fixed model.

Results

The results of this analysis indicated no significant differences for either of the methods, for either of the procedures or for their interaction. However, there were several results which approached significance—two isolated effects and two rather important effects.

The isolated effects were found on the first posttest. On Part B, which was used to measure immediate transfer, an F value of .07 resulted for the interaction of method and procedure. Examination of the means would seem to indicate that the better combinations were the didactic method with the independent procedure, and the discovery method with the teacher-guided group procedure.

Part C of the first posttest, which measured immediate problem solving, yielded an F value of

.07 for procedure with higher means for the two groups using the independent procedure. This result was probably due to task similarity since the two independent groups used work sheets similar to the posttest sheets to learn the principle in the lessons. By the time of the second posttest, this difference had ceased to exist.

The two important results were found on the second posttest. Results which approached significance were found on Part B, the delayed transfer test, for both of the main effects. A probability level of .09 resulted for method with a level of .08 for procedure. Examination of the means pointed to the discovery approach as the better method and to the teacher-guided group technique as the better procedure.

Discussion

The failure to find very significant results in this study may have been due to several factors. One of the basic concerns had been to work within the context of the regular school program and to avoid any reactive effects due to special experimental procedures. Consequently, classes were kept intact, the regular classroom teachers carried out the program, and the experimental treatment was part of a total program. These arrangements may have caused several problems: classes may have had differential exposure to the spelling principles between the time of the pretest and the program; teachers may have lacked skill in handling the discovery method because they received no special training; and—most important of all—the amount of practice and review in the lessons may have been sufficient to override the effects of the differential treatment used to teach the principle.

Two other factors may also have been responsible for the results. Retention was tested at the end of the 6-week program rather than immediately following the experimental treatment—a fact which may account for the lack of support found for the didactic method on the retention test. However, several of the procedures which had been instituted to give the didactic treatment a fairer trial should have been of benefit to it. These procedures included establishment of the same goals and use of the same principles as for the discovery treatment, as well as the use of a great deal of repetition in teaching the principle in order to equate the amount of time used for both methods.

Conclusions

In view of the problems involved in attempting to simulate the normal school situation, it is hardly surprising that the results failed to meet the usual criteria for statistical significance. It is surprising, however, that the didactic method failed to show any advantage on the retention test in light of the fact that the discovery method did show some of its usual advantage on the delayed transfer test. In conclusion, if one is willing to consider the results which approach significance, the findings of this research study would seem to indicate a tentative place for the discovery method and for the teacher-guided group procedure in the sphere of principle learning—for the spelling curriculum in particular and perhaps for other school curricula as well.

Implications for Further Research

The goal of this research had been to avoid the many pitfalls noted by Cronbach and others which abound in so many of the past studies on discovery learning. Because of this concern, methods were used which would not handicap the didactic groups and which would avoid student awareness of the experimental nature of the program. The learning task was part of educational subject matter and the program was long enough to avoid the effect of one or two isolated sessions.

From the discussion, it is obvious that these practices may have become liabilities. Consequently, it may be necessary to sacrifice some of these practices in order to obtain significant results. If educational subject matter is to be used, the content should probably be related to but outside of the mainstream of the regular curriculum. If the experiment is to be carried out in the regular classrooms, teachers should be given special training. A pretest based on the experimental program should be given, and students should be randomly assigned by the experimenter in order to reassure equality of groups—even though the results may reflect reactions to these experimental practices. The results of this study would seem to suggest acceptance of some of these compromises in order to achieve statistically significant results.

REFERENCES

1. Ausubel, David P., *The Psychology of Meaningful Verbal Learning,* Grune and Stratton, New York, 1963.
2. Bruner, Jerome S., "The Act of Discovery," *Harvard Educational Review,* 31: (no.1) 21–32, 1961.
3. Cronbach, Lee J., "The Logic of Experiments on Discovery," in Shulman, L. S.; Keislar, E. R. (eds.), *Learning by Discovery: A Critical Appraisal,* Rand McNally, Chicago, 1966, pp. 77–92.
4. Hildreth, Gertrude, *Teaching Spelling,* Henry Holt, New York, 1955.
5. Horn, E., *Teaching Spelling,* Department of Classroom Teachers of the American Research Association of the National Education Association, Washington, D. C., 1954.
6. Wittrock, M. C., "The Learning by Discovery Hypothesis," in Shulman, L. S.; Keislar, E. R. (eds.), *Learning by Discovery: A Critical Appraisal,* Rand McNally, Chicago, 1966, pp. 33–76.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

A 5-Year Follow-Up Study of High Ability Achieving and Nonachieving College Freshmen

LELAND SMITH, University of Kentucky

ABSTRACT

This study investigated the present status and significant differences between high ability achieving and non-achieving college freshmen of a 1960–61 study. The Ss were students at the University of Kentucky and had scored in a percentile range of 95 to 99 on the College Qualification Tests (CQT). The psychosocial data were collected by means of a questionnaire, and significant differences were determined through the chi-square method of statistical analysis. The resulting profiles indicate the possibility that many high ability non-achievers might achieve if they were given early guidance and were allowed the opportunity to change their academic programs.

ANY UNIVERSITY that chooses to do so can make a major reduction in the proportion of its students who drop out, unless that proportion is already quite small (2). This is especially true when dealing with students who rank in the top 5 percent in academic ability. It is estimated that 20 percent of the most able students admitted to college withdraw or are dropped before graduation, resulting in a setback in their academic careers (5). In a 1966 study of high ability freshmen, Seibel found that after 1 year of college 9 percent of the males were on academic probation; 6 percent had withdrawn for academic reasons, and 3 percent had withdrawn for non-academic reasons (3). These students appear to be victims of inadequate counseling, the wrong curricula, and various pressures of the multiversity.

It could be that many of the wrong people were dropped, for most of these students eventually recover and finish their degrees. For example, Eckland stated that there may be far fewer students who permanently give up their college careers than previously thought, as most studies fail to make allowances for the prolonged nature of academic careers and the dropouts who come back (1). Is it accurate to label these students as failures and dropouts? Are there significant differences in the status of those designated as achievers and those designated as non-achievers? Have these high ability non-achievers been done a disservice by dropping them from the university environment thus making it difficult for them to make an academic adjustment? Answers to some of these questions may be implied from the present profiles of some achievers and non-achievers of 5 years ago.

Background

The present study is a 5-year follow-up of high ability achieving and non-achieving college freshmen. In the fall of 1960, 154 male freshmen at the University of Kentucky scored in the top 5 percent on the CQT. It was assumed that a percentile score range of 95 to 99 on the CQT was indicative of ability to do very good college work. However, thirty-eight of these high-ability students failed to achieve a GPA (grade point average) of 2.00 (a standing of "C") in their academic work for the first semester. At the close of the second semester, a total of sixty-two of these high-ability freshmen had failed to achieve a standing of 2.00 during either the first or second semester or both. Thus 40 percent of the potentially most capable freshmen failed to achieve acceptable grades during their first year at the University. It was concluded that the significant difference between the 25 percent non-achievers and the 26 percent achievers was of a noncognitive or psychosocial nature, rather than any lack of ability to do college work. It should be noted that forty of the total group of 154 in the top 5 percent of freshmen did achieve a GPA of 3.00 (a standing of "B") or better.

The purpose of the 1960 study was to identify significant psychosocial differences between achievers and non-achievers as revealed by interview data. The interview included sixty-seven items relative to the students' psychosocial backgrounds, and was administered to thirty-one non-achievers and thirty-two achievers of high academic potential. The significant results of the 1960 study were published in *The Journal of Educational Research* (4).

Procedure and Results

In the fall of 1966, a 50-item questionnaire was mailed to each of the original thirty-one non-achievers and thirty-two achievers. Twenty-two non-achievers and twenty-eight achievers responded. The questionnaire covered the following areas: (a) present academic status, (b) socioeconomic status, (c) perceptions of the pre-university experience, (d) perceptions of the university experience, and (e) present values, aspirations, and pressures.

The significant differences of attitude, opinion, and perception between the achievers and non-achievers were determined by the chi-square method of statistical analysis. For differences to be significant, a probability of less than .05 was established as a standard.

The following findings were based on the significant differences of attitude and opinion obtained from the questionnaire data. Approximate levels of probability may be obtained by referring to Table 1.

Present Academic Status. After the passage of 5 years, the achievers were significantly more likely to have obtained a college degree than the non-achievers were. The difference was significant.

More achievers than non-achievers are still enrolled in some capacity, medical school, or graduate school. Of the twenty-eight achievers responding, four are candidates for the PhD, three are senior medical students, six have earned the MA degree, one has a law degree, twelve hold the BA or BS degree, and two are undergraduates. The difference was not significant.

A change of major was considerably more frequent among non-achievers than achievers. Only 19 percent of the achievers had changed their majors, while 63 percent of the non-achievers had changed their programs. The difference was significant.

A summary of the educational status of the twenty-two non-achievers is as follows: one is a PhD candidate, four have earned the MA degree, one is pursuing the MA, and four hold undergraduate degrees. Seven are still actively pursuing undergraduate degrees, and five are no longer enrolled in a degree program.

Table 1.—Chi-square Results for Significance of Difference Between Freshman Achievers and Non-achievers

	df	x^2	P
PRESENT ACADEMIC STATUS			
1. College Degree	1	13.966 *	< .001
2. Presently Enrolled	1	2.748 NS	< .10
3. Change of Major	1	9.165 *	< .01
SOCIO-ECONOMIC STATUS			
1. Satisfaction with Income	1	6.165 *	< .02
2. Non-College Training	1	6.881 *	< .01
PRE-UNIVERSITY EXPERIENCE			
1. Satisfaction with H. S. Preparation	2	5.326 NS	> .05
2. High School Guidance Service	1	4.967 *	< .05
3. Quality of H. S. Teaching	1	12.913 *	< .001
4. Parental Guidance	1	0.583 NS	< .50
THE UNIVERSITY EXPERIENCE			
1. Sense of Intellectual Growth	1	8.961 *	< .01
2. Professors are Interested in Student Adjustment	1	3.384 NS	> .05
3. Expectations of the College Experience	1	2.204 NS	> .10
4. Difficulty of Friendship with Professors	1	4.970 *	< .05
5. Regret of Non-achievement	1	4.002 *	< .05
6. Evaluating College Teaching	1	0.866 NS	< .50
7. Would Choose the University Again	1	3.447 NS	< .10
VALUES, ASPIRATIONS AND PRESSURES			
1. Shift in Value System	1	4.419 *	< .05
2. Degree of Value Change	1	3.271 NS	< .10
3. Aspirations	1	0.398 NS	< .70

* Significant at .05 level. NS Not Significant

Socioeconomic Status. The non-achievers were significantly more dissatisfied with their income in relation to their perceived ability and potential than were the achievers. This was true even though 64 percent of the achievers were still in the process of pursuing advanced degrees. The difference was significant.

More non-achievers had taken non-college training, such as business school, sales training, and on-the-job training, than had the achievers. This trend was in the expected direction and was significant.

Perceptions of the Pre-University Experience. It was found that there was generally greater satisfaction among achievers than non-achievers with the quality of their high school preparation. Though not significant, this difference was in the direction one might expect.

Non-achievers rated high school guidance services as poor; achievers, on the other hand, felt their guidance had been satisfactory. The difference was significant.

Over 28 percent of the non-achievers felt that the quality of high school teaching was poor, while all the achievers rated high school teaching as excellent or satisfactory. The difference was significant.

Eighty-five percent of the achievers and 81 percent of the non-achievers perceived that the guidance they had received from their parents had been realistic. The difference was not significant.

It is worth noting that 92 percent of the achievers and 82 percent of the non-achievers felt that their parents had not brought undue pressure to bear in the selection of their academic majors.

Perceptions of University Experience. All of the achievers felt that they experienced intellectual growth during their first year at the University; 28 percent of the non-achievers felt that the first year was an academic loss. The difference was significant.

It should be noted that achievers tended to feel more frequently that professors are interested in student adjustment. Non-achievers felt more frequently that professors are not interested in student adjustment. The difference was not significant.

In their expectations of the college experience, achievers appear to have been more realistic than non-achievers. Although not significant, differences were in the expected direction.

Non-achievers perceived that it was difficult to become friendly with professors more frequently than did achievers. The difference was significant.

Significantly more non-achievers regretted a failure to make the most of their educational opportunities while at the University.

In retrospect, both achievers and non-achievers rated teaching at the University as satisfactory. Seventy-five percent of the achievers and 86 percent of the non-achievers rated the instruction positively. The difference was not significant.

Three-fourths of the total group responded "yes" that they would choose the University again. Although more achievers than non-achievers said they would and more non-achievers said they would not choose the University again, the difference was not significant.

The non-achievers of this study who had been dropped were unanimous in the opinion that this disruption of their academic careers had hindered rather than helped in the development of their talents.

Values, Aspirations and Pressures. In general, both achievers and non-achievers felt that their values had shifted over the past 5 years. Seventy percent of the achievers and 75 percent of the non-achievers considered that their ideas of what is important in life had changed. Achievers indicated their values had shifted from an idealistic to a more realistic and "cynical" attitude toward the world. Non-achievers on the other hand shifted from goals of personal pleasure to those of deeper concern for responsibility and self-realization. The difference was significant.

Non-achievers felt that their values had changed a great deal, while achievers said they had changed to a lesser degree. The difference was not significant.

In the area of aspirations, achievers and non-achievers were quite similar. Both groups were much more concerned with personal goals than they were with making contributions to society or being of service to humanity. Sixty-four percent of the achievers were most concerned with personal happiness and 32 percent felt their vocations were of most import. Most achievers aspired for personal happiness and success. (Perhaps the aspiration for success is indicative of a greater need to prove one's self, especially after having experienced academic failure and probation.) The difference was not significant.

The achievers and non-achievers described similar psychosocial problems and pressures. The achievers most often mentioned draft status, monetary, and academic pressures as being most urgently felt. The non-achievers were different in that they expressed an almost driven concern to make better than average grades and to finish their degrees. This attitude could be due to their initial underachievement and the onus of probation. The difference was not significant.

Conclusions

The responses to questionnaire items, by the twenty-eight achievers of high ability and twenty-two non-achievers of high ability in a 5-year follow-up study, have led the investigator to the following conclusions which are relative to the specific groups studied.

According to this study, there was a trend for achievers to attain a degree of self-definition and commitment earlier in life than non-achievers.

Neither marriage nor date of marriage appears to be a significant factor in achievement or non-achievement as approximately half of both groups were married.

The psychological impact of being placed on academic probation, especially to bright students, appears to be considerable. Seventy-two percent of the non-achievers said that the greatest impact was to their own self-esteem, while 28 percent felt that their parents were affected to a greater degree.

The terms failure and dropout, when applied to those designated as non-achievers, are misnomers for it was found that 75 percent of these students achieved on their return to an academic community.

The dropping of high ability students is not the answer to non-acihevement. Rather, the question should be asked, "What is wrong?" Many high-ability non-achievers in our universities might succeed, if guidance were given early in the freshman year, and if university policy allowed students to change their commitments instead of dropping them.

Limitations

It should be noted that these findings are not to be interpreted as definite proof that a causal relationship exists between the variables under investigation and college failure. However, these retrospective results do lead to hypotheses which may be investigated with suitable controls, making predictions in advance of failure. For the present, the conclusions of this study are primarily suggestive in nature.

REFERENCES

1. Eckland, Bruce K., "College Dropouts Who Came Back," *Harvard Educational Review*, Summer 1964, p. 403.
2. Ford, Donald H.; Urban, Hugh B., "College Dropouts: Successes or Failures," *The Educational Record*, Spring 1965, p. 78.
3. Seibel, Dean W., "The Relationship of Some Academic Ability Characteristics of High School Seniors to College Attendance and Performance," *College and University*, Fall 1966, p. 47.
4. Smith, Leland, "Significant Differences Between High Ability Achieving and Non-achieving College Freshmen as Revealed by Interview Data," *The Journal of Educational Research*, September 1965, pp. 10–12.
5. Wolfle, Dael, *America's Resources of Specialized Talent*, Harper and Row, New York, 1954, p. 150.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

The Effects of a Mediational Device on Paired–Associate Learning; Learning and Retention of Sixth Grade Students

WILLIAM E. LOADMAN, Indiana University

ABSTRACT

The objective of this experiment was to determine the effect of different levels of mediational aids on a paired associate task when presented to sixth-grade students for amount and rate of original learning and long term retention. Three levels of strategies were employed. High level mediational aids were found to significantly effect the amount and rate of original learning; the retention data also suggest similar inferences.

THE PURPOSE of this investigation was to determine the influence of different levels of strategies for paired-associate (PA) learning on accuracy, speed of learning, and long-range retention. Many PA studies have been conducted, but only recently has attention been given to mediational aids and the long-term retention of the learning task.

Jensen and Rohwer (3) using a PA task demonstrated that invoking a mnemonic device on the learning task facilitates immediate recall. However, this study cannot be generalized beyond mentally retarded subjects. In a follow-up experiment, Jensen and Rohwer (2) conducted a 1-week delayed retention task with retarded Ss using mnemonic devices on the original learning task. They found that the group receiving the mnemonic aid responds slightly but not significantly better than a control group. Davidson (1) found that the mnemonic device facilitates learning and that high ability students are better able to use mediators than low ability students.

Martin and others (4) using students on a PA task discovered that in learning the task, the Ss employed different levels of mediational strategies and that those Ss using the higher level strategies responded correctly significantly more often than those using lower level strategies. This study also identified and classified the different levels of strategies. Further, normal and mentally retarded children were compared in their use of learning strategies on a PA task. The results indicated that normal children used higher level strategies than retarded children, and also learned the list faster and more accurately than the slower children. Using fourth-, sixth-, and eighth-grade Ss, they found that the use of high level strategies increases with age and results in a greater number of correct responses. A 48-hour retention task was conducted and the results indicate that the retention scores were directly related to the level of strategy used by the S, i.e., students employing high level strategies scored high on the retention task. The results indicate that strategies aid the learner on a delayed retention task which is in opposition to those found by Jensen and Rohwer (2). Martin and others (5) also found that Ss recall more words correctly after they have been given high level strategies on half of the PA task than a control group; the experimental group also used higher level strategies, on those dyads without strategies provided, than the control group.

This research suggests that a mediational device facilitates learning and that some Ss are more able to employ higher level strategies than others. It is not clear exactly what effect the mediational device has upon long term retention. Therefore, is it possible to give all students high-level mediational devices to facilitate learning, and if so, what effect do these aids have upon long-term retention? Also, is it better for Ss to derive their own high level strategies or have the high level strategies given to them?

Procedure

The Ss consisted of three randomly selected groups of sixth-grade children from a Midwestern Catholic school in a middle-class neighborhood. After the students were randomly assigned to a group, the groups were randomly selected for a control and two experimental conditions. A total of seventy-six students participated in the experiment and all were naive with respect to PA learning tasks.

Thirteen PA dyads were constructed from Noble's (6) list and consisted of low-meaningful (L-M) stimulus disyllables and high-meaningful (H-M) response disyllables. Five of these pairs were used in a practice task and the remaining

Table 1.—The Means and Standard Deviations on Practice, Learning, Trials to Criterion and Retention of the Paired-Associate Task[1]

		Control	Sentence Given	Derived Mnemonic
Number	N=	25	25	24
Correct on	$\bar{X}$=	5.36	7.36	6.5
Practice Task	S=	3.21	3.47	3.88
Learning Trials	N=	26	25	25
1-7 Number	$\bar{X}$=	36.31	46.64	30.30
Correct	S=	12.31	4.29	16.47
Learning Task	N=	26	25	25
Trials to	$\bar{X}$=	6.27	4.44	6.76
Criterion	S=	1.61	1.39	1.71
Retention Task	N=	26	23	25
Number Correct	$\bar{X}$=	6.46	7.43	4.92
on Trial 1	S=	2.32	1.16	3.23

[1] The inconsistency in the number of Ss responding was caused by a lack of understanding on the practice task and mortality on the retention task.

eight were used in the experiment. Obvious association on the pairing of the words during construction of the dyads was held to a minimum.

Each group was given a practice task, a learning task, and a 1-week retention task. All groups were given the same instructions and directions on the practice task and each group was presented the same material. Each learning trial was followed by a recall trial which allowed the S to write on the answer sheet what was thought to be the correct H-M response to the L-M stimuli. A learning trial was considered to be the individual presentation of the five PA's. A recall trial was considered to be the individual presentation of the five stimulus words. The learning trials were presented three times and each learning trial was spaced by a recall trial. This allowed the Ss to familiarize themselves with the learning task and also allowed the experimenter (E) to check for comparability of groups. A recall task was used in all phases of this experiment.

After the practice task, the Ss were given instructions and directions according to the specific experimental treatment group.

Control Group. The control group was told that they would now be given a similar task to learn. They were given no special aids for learning.

Sentence Given Group. This group was told that it was more efficient to learn the task if the words are in sentences. Therefore, a sentence with the PA words in it was read aloud to the group as each PA was presented in learning trials one and two. An example was given to demonstrate this procedure.

Derived Sentence Group. This group was told that it was more efficient to learn the task if the words are in sentences or if some association is made. A few examples were given to demonstrate this procedure and the group was reminded to put the PA's in sentences on learning trials one and two.

On the experimental task the PA's were individually presented on the learning trials and the stimuli were presented on the response trials; both were presented in random order. The criterion task consisted of eight PA's with eight learning and eight recall trials. The first seven learning and recall trials were similar to the trials on the practice task. However, on the eighth recall trial the time presentation was decreased on the stimulus; the learning trial remained the same. It was reasoned that if one method of learning is superior to another in terms of the acquisition of PA's, then the better one knows the association, the faster he can respond when one member of the pair is presented. If the relationship were not spontaneous, then the S could not respond fast enough to complete the task.

One week after the original learning occurred, the E returned to the school for a long term retention task. After the directions, the Ss were given a recall task exactly like the trials employed in the original response trials. The eight stimulus words were presented individually on trial one and the Ss responded accordingly.

Results and Discussion

On the practice task, the three treatment groups were compared for the number of correct responses over the three trials by a one-way analysis of variance. These groups were not significantly different at $P < .05$ level on this task ($F = 1.99$; $df = 2,71$; see Tables 1 and 2). One notices that the sentence given group (SG) has a higher mean than either the derived sentence (DS) or control (Con) group. The Con group is particularly low, but these differences are nullified by the large variance within each group. The instrument used was rather short and crude and may not have

Table 2.—One Way Analysis of Variance on Practice, Learning, Trials to Criterion and Retention of the Paired Associate Task

	Source	df	MS	F
Number Correct on Practice Task	among	2	25.16	1.99
	within	71	12.6692	
Learning Trials 1-7 Number Correct	among	2	1765.1929	11.90**
	within	73	148.3465	
Learning Task Trials to Criterion	among	2	37.4705	15.04**
	within	73	2.4905	
Retention Task Number Correct on Trial 1	among	2	38.8339	7.20**
	within	71	5.3966	

** Significant Difference at .01 level
F.99(2,70)=4.92

given reliable or valid estimates of performance. However, there is no reason to suspect that this is true, and it appears logical to conclude that the samples used in this experiment are comparable.

The criterion task was scored for the number of correct responses on trials one through seven. A one-way analysis of variance was performed on these data, and the differences were found to be significant at the $P < .01$ level ($F = 11.90$; $df = 2{,}73$; see Tables 1 and 2). A Scheffè Test of Multiple Comparisons was performed on the mean scores and the SG group was statistically superior to both the Con group and the DS group. The Con and DS groups did not differ significantly.

On the learning task, the SG group was significantly superior in reaching a criterion of two perfect response trials than either the DS group or the Con group according to a one-way analysis of variance and the Scheffè Test of Multiple Comparisons ($F = 15.04$; $df = 2{,}73$; $P < .01$ see Tables 1 and 2). However, the DS and Con groups were not statistically different. This suggests that invoking a high level strategy facilitates the rate of learning. The Con and DS groups did not differ because the Con group had to learn the task on their own, and the DS group needed learning time to develop their own high level strategies which didn't allow them to learn the list as quickly as the SG group. In a PA learning context it appears most beneficial as far as speed of learning and number of correct responses are concerned to give Ss the high level strategy. A limited amount of generalization to a school learning context seems warranted if one can equate a PA task with a particular learning task. This being the case, it appears to be most helpful for the student to be given a high level strategy as opposed to having the student derive his own high level strategy or be left to his own resources. This result appears to be at least valid at the sixth-grade level. For example, if one were attempting to teach the vocabulary of a foreign language, one might put the foreign word and its English equivalent into a short sentence or phrase to make an association for the student.

Recall trial eight will not be considered in this paper. On the retention task, one could legitimately argue that the amount of retention is not independent of original learning. However, all Ss received the same number of trial presentations for the same amount of time and the majority of Ss reached the criterion on the learning trials. Therefore, the E considers the retention task to be an independent measure. The number of correct responses per group were analyzed by means of a one-way analysis of variance and a Scheffè Test of Multiple Comparisons. The SG group was statistically superior, in the number of words correctly recalled, to the DS group but not different from the Con group ($F = 7.20$; $df = 2{,}71$; $P < .01$; see Tables 1 and 2). Unexpectedly, the Con group and the DS group were not statistically different. This may be explained by the difference in amount and rate of original learning. However, a more plausible explanation seems to emerge when one examines the individual effects of the treatment groups. The scores of the SG group appear to be relatively smooth across individuals in both learning and retention thus producing a low variance. The Con group scores are less stable across individuals and not as high overall as the SG scores are and thus have a greater amount of variance. The DS group scores are very unstable across individuals and this not only produces a relatively low mean, but also a very high variance. The original learning data and even more so for the retention data seem to suggest that telling the Ss that a high level strategy is most beneficial for learning and then having them derive their own devices does not produce a unitary effect. Instead, two extreme effects appear to be prominent; one type of S is able to develop this strategy and compares favorably with the SG group on retention; the second type seems to have difficulty invoking the high level strategy on his own and performs relatively poorly on the learning and retention task. This suggests that a great deal of structure is more beneficial than a limited amount in a PA learning task for sixth-grade Ss. This has serious implications for any type of discovery learning that can be equated with a PA task. It suggests that while some students can learn and retain associations by a limited structure technique (learning and retention equal to that of a high-level structure), it is more beneficial to impose high level structure on the task for better overall performance. This is at least true at the sixth-grade level under the conditions and restrictions generated in this study.

REFERENCES

1. Davidson, R. E., "Mediation Ability in Paired-Associate Learning," *Journal of Educational Psychology,* 55:352–356, 1964.

2. Jensen, A. R.; Rohwer, W., "The Effect of Verbal Mediation on the Learning and Retention of Paired-Associates by Retarded Adults," *American Journal of Mental Deficiency,* 68:80–84, 1963.

3. Jensen, A. R.; Rohwer, W., "Verbal Mediation in Paired-Associate and Serial Learning," *Journal of Verbal Learning and Verbal Behavior,* 1:346–352, 1963.

4. Martin, C. J.; Boersma, F. J.; Cox, D. L., "A Classification of Strategies in Paired-Associate Learning," *Psychonomic Science,* 3:455–456, November, 1965.

5. Martin, C. J.; Cox, D.; Bulgarella, R., "Verbalization of Associative Strategies at Three Developmental Levels," paper prepared for the 1966 APA Convention, New York.

6. Noble, C. E., "An Analysis of Meaning," *Psychological Review,* 59: 421–430, 1952.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Control of Reinforcement and Academic Achievement[1]

RALPH L. SHAW and NORMAN P. UHL, Emory University

ABSTRACT

The purpose of this study was to investigate the relationship between children's locus of control scores and school achievement, as measured by reading scores, for low and upper-middle socioeconomic level (SEL) white and black groups. It was found that the low SEL group had significantly higher external scores than the upper-middle SEL group and that within the upper-middle group the blacks had higher external scores than the whites. However, within the low group, there was no relationship between race and external scores. Of the four groups, white low SEL, black low SEL, black upper-middle SEL, and white upper-middle SEL, the locus of control scores related to success in reading in only the latter group. It was found that the higher the external score, the lower the reading score. The findings are discussed in terms of their educational implications.

THE PERFORMANCE of the child in the classroom has been of great concern to educators and psychologists for many years. Many studies examining the relationship of IQ, sex, race, and socioeconomic level variables to classroom achievement have been performed. With increased emphasis on individualized instruction, the consideration of personality factors has gained importance. This study is directed toward investigating the relationship between academic achievement, as measured by reading scores, and one personality variable: locus of control of reinforcement.

Locus of control of reinforcement can be of two types: internal control and external control. Internal control of reinforcement refers to the belief that the reinforcements a person receives are a result of his own purposeful action. If the individual believes that the reinforcements he receives are due to some force beyond his control, such as luck, or fate, then he is said to have an external locus of control.

Lefcourt (6), in his review of the literature, relates this personality construct to the concepts of competence and effectance, helplessness and hopelessness, striving toward mastery, and alienation. The social learning theory of Rotter (8) looks upon this construct as a generalized expectancy concerning reinforcements.

Social class and ethnic membership have been shown to be important variables in determining the locus of control of reinforcement. Several studies, such as those of Battle and Rotter (1) and Lefcourt and Ladwig (7), found that blacks score higher on external control than whites and that lower-class individuals express greater externality than middle-class persons.

The relationship of IQ to locus of control of reinforcement is not as clear. Studies by Bialer (2), using elementary school subjects, and Crandall, Katkovsky, and Crandall (4), using elementary and high school students, have shown a positive relation between IQ and locus of control, with higher internal control scores being associated with higher IQ levels. Battle and Rotter (1) found that black, 11-13-year-old students tended to have higher external scores regardless of IQ status, when compared to white students.

Conflicting evidence has been found regarding the relation of sex to locus of control of reinforcement. Girls were shown to be more likely to give responses indicating internal control than were boys in one study (4), while no relation between locus of control and sex was found in a study by Battle and Rotter (1). Few studies have investigated the relationship between locus of control scores and school achievement. Crandall, Katkovsky, and Crandall (4) found that their scale predicted differently for the two sexes at different age levels.

This study examined the effect of SEL, (low and upper-middle), race (black and white), and sex of second graders on locus of control scores and also investigated the relationship between locus of control scores and reading achievement. It was predicted that: (1) low SEL Ss have higher external scores than upper-middle SEL Ss; (2) blacks have higher external scores than whites; (3) there is no sex difference on locus of control scores; (4) locus of control scores are inversely related to reading scores, the higher the external scores, the poorer the reading scores; and (5) locus of control scores are not related to intelligence scores.

Method

Subjects. The Ss in this study, 211 second-grade public school children from a large metropolitan area, participated in a larger study. The Ss were drawn from eighteen classrooms in six elementary schools. Three of the schools were in upper-middle SEL areas and three were in low SEL areas. In each SEL area, two of the schools were predominantly white and one was predominantly black. Careful attention was given to the selection of the schools in order to obtain homogeneity of SEL within each school.

A stratified random sampling technique was employed to insure adequate numbers of males and females, blacks and whites, and low and upper-middle SEL Ss. The mean age at the time of testing was 8.0 years, with a SD of .66 years. The sample was composed of 104 low SEL and 107 upper-middle SEL Ss, 69 blacks and 142 whites, and 107 males and 104 females.

Instruments. The paragraph meaning and word reading sections of the Primary I and Primary II levels of the 1964 edition of the Stanford Achievement Test, Form W, were administered to obtain a measure of reading achievement. Four subtests from Cattell's Culture Fair Intelligence Test, Scale I, Substitution, Classification, Mazes, and Similarities, were used to measure IQ. These scales were used for practical reasons, in addition to Cattell's (3) statement that the use of these subtests results in a more fully culture fair IQ.

The Bialer-Cromwell Children's Locus of Control Scale was used to measure degree of internal-external control of the Ss. This test consists of twenty-three items worded so that a "yes" or a "no" response is required for each item. A "yes" response indicates internal control on some items and external control on other items. The sum of the external responses equals the external control score.

Procedure. All tests were administered within a 2-month period. The Stanford Achievement Tests were group administered according to the instructions contained in the Manual. The Classification subtest of Cattell's Culture Fair Intelligence Test of the Bialer-Cromwell Children's Locus of Control Scale were individually administered according to the instructions in their Manuals, each taking approximately 10 and 4 minutes, respectively, to administer.

The remaining three subtests of the Culture Fair Intelligence Test (Substitution, Mazes, and Similarities) were administered to groups of no more than fifteen students using the instructions given in the Manual.

Results

A Kuder-Richardson reliability of .95 was obtained for the reading score. Test-retest reliability of .85 was obtained with the Culture Fair Intelligence Test, with approximately 1 week separation between tests. A Kuder-Richardson reliability of .43 was found for the Children's Locus of Control Scale.

An analysis of variance using a 2 × 2 × 2 factorial design was employed to investigate the effect of SEL, race, and sex on locus of control scores. The main effect of SEL was found to be significant [$F(1,204) = 3.95, p < .05$] in addition to the interaction of SEL and race [$F(1,204) = 4.21, p < .05$]. No other main effects or interactions were significant at the .05 level.

The significant SEL × race interaction was investigated by comparing the means of the black low SEL group, the white low SEL group, the back upper-middle SEL group, and the white upper-middle SEL group which were 12.19, 12.29, 12.16, and 11.05, respectively. The respective standard deviations were 2.90, 2.52, 2.18, and 2.79. Through the use of Scheffé's Test (9) it was found at the .05 level of significance that the white upper-middle SEL group had lower external scores than the other three groups and that there was no difference in the means of these latter three groups. Thus, the predictions that low SEL Ss have higher external scores than upper-middle SEL Ss and that there are no sex differences are supported. The prediction that blacks have higher external scores than whites is only true of the upper-middle SEL Ss.

In only one of the four groups studied was there a significant relationship between locus of control and success in reading. In the white upper-middle SEL group it was found that the higher the external score obtained by the student, the lower would be his score in reading ($r = -.31, p < .01$). In the other three groups, however, the correlations ranged from .09 to .17 in magnitude, all nonsignificant at the .05 level. IQ was related to reading in both SEL groups ($r = .47$ and .23 for upper-middle and low SEL, respectively), but IQ was not related to external scores. When reading and external scores were correlated with IQ partialed out, no change in the relationships was observed.

Discussion

The low reliability of .43 obtained using the Children's Locus of Control Scale was a surprise. This scale in one of the two most frequently used instruments to measure internality–externality in children. In reviewing the published studies which have made use of this instrument, the authors could not find one which reported reliability. However, in a study performed concurrently with this one, Shaffer, Strickland, and Uhl (10) report a reliability of .49 for this scale using fourth-grade Ss from two SEL's comparable to the one used in the present study. These results indicate need for further work on this scale.

Results support predictions 1, 3, and 5, namely, low SEL Ss were found to have higher external scores than upper-middle SEL Ss, no sex differences on locus of control scores were found, and locus of control scores were not related to intelligence scores. Predictions 2 and 4 were partially supported. It was found that blacks have higher external scores than whites, but only in the upper-middle SEL group. It was also found that locus control scores were inversely related to reading scores, but only in the white upper-middle SEL group.

The finding that low SEL Ss have higher external scores agrees with the results of other studies (1, 4). The low SEL groups in the present study and in the studies cited were found to score significantly higher on external control than the middle-class groups. Battle and Rotter (1) maintain that there is a lesser belief in self-responsibility among lower-class Ss than is found in the upper-class Ss, while others (4) state that there is less opportunity to manipulate the environment successfully due to lack of education, money, and status among the lower class groups. Lefcourt (6:212) says, "Perhaps the apathy and what is often described as lower-class lack of motivation may be explained as a result of the disbelief that effort pays off."

Battle and Rotter (1) found that low SEL blacks rated highest on external control of the four groups studied, while the middle-class group rated highest on internal control. This relationship between race and external scores was not found among the low SEL groups, but there was a significant relation found between these two variables among the upper-middle SEL groups, with the blacks more external than the whites.

Similar to Battle and Rotter's (1) findings, but contrary to Crandall, Katkovsky, nd Crandall's (4), there was no relation between IQ or sex and locus of control found in this study. The finding in this study that locus of control was significantly related to success in reading among only one of the groups studied (the white upper-middle SEL group) is one which needs to be examined further. It is hypothesized that in the upper-middle SEL home and in the school the importance of reading is stressed repeatedly to the young child. It is very early put into their value system with a high positive valence. Lefcourt (5) postulates that Ss with internal locus of control are able to see reward contingencies and to react to them in an appropriate manner, while external Ss fail to recognize the reward contingencies available to them. Thus, the child with an external locus of control may well have a high value placed on success in reading, but may not be able to perceive reward contingencies and thus may fail to achieve success in reading.

In the other three groups studied, reading may not be stressed so highly in the home, even though it may be stressed in the school, so that Ss fail to incorporate the importance of reading into their value system. With no particular value assigned to the importance of success in reading, whether the S is internal or external in relation to his locus of control would have little bearing on his achievement, which is what was found in this study.

If this interpretation of the findings is correct, it suggests the importance of educating parents to the importance of the values of the school. With more active support for the school's goals from the parents, the positive incorporation of the school's goals into the value system of the child should be facilitated.

FOOTNOTE

1. We express our gratitude to Mrs. Annette Lashner who assisted in collecting part of the data and to the Urban Laboratory in Education (Atlanta Education Improvement Project) for financial support of this study. The computer time was supported in part through the facilities of the Emory University Computer Center.

REFERENCES

1. Battle, E.; Rotter, J. B., "Children's Feelings of Personal Control as Related to Social Class and Ethnic Group," *Journal of Personality*, 31: 482–489, 1963.
2. Bialer, I., "Conceptualization of Success and Failure in Mentally Retarded and Normal Children," *Journal of Personality*, 29: 303–320, 1961.
3. Cattell, R. B., *Culture Fair (or Free) Intelligence Test*, Institute for Personality and Ability Testing, Champaign, Illinois, 1950.
4. Crandall, V. C.; Katkovsky, W.; Crandall, V. J., "Children's Beliefs in Their Own Control of Reinforcements in Intellectual-Academic Achievement Situations," *Child Development*, 36: 91–109, 1965.
5. Lefcourt, H. M., "Belief in Personal Control: Research and Implications," *Journal of Individual Psychology*, 22: 185–195, 1966a.
6. Lefcourt, H. M., "Internal Versus External Control of Reinforcement: A Review," *Psychological Bulletin*, 65: 206–220, 1966b.
7. Lefcourt, H. M.; Ladwig, G. W., "The American Negro: A Problem in Expectancies," *Journal of Personality and Social Psychology*, 1: 377–380, 1965.
8. Rotter, J. B., *Social Learning and Clinical Psychology*, Prentice-Hall, Englewood Cliffs, New Jersey, 1954.
9. Scheffé, H., "A Method for Judging All Contrasts in the Analysis of Variance," *Biometrika*, 40: 87–104, 1953.
10. Shaffer, Scott; Strickland, Bonnie R.; Uhl, Norman P., "The Relationship of Individual Difference Measures to Socioeconomic Level and to Discrimination Learning," paper presented at the annual convention of the Southeastern Psychological Association, New Orleans, February 1969.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Differentiating Motivations Underlying Vocational Choice

DANIEL COHEN, Pace College, New York

ABSTRACT

This study was designed to determine whether or not expressed motivations for curricular choice would discriminate between college seniors in business administration and in teacher education. A list was compiled of all motivations expressed by 504 male students, selected randomly, from two colleges in New York City. After eliminating duplications, 289 items were selected, weighted, and combined into a questionnaire with a 5-point intensity scale. The t ratio and factor analysis were used to isolate the common underlying "teacherness" and "businessness" within groups. Major significant factors ($P < .001$) were then programmed into the discriminant analysis. The factor weights of two important factors (indicated by F ratios) were used as the coefficients for the final discriminant equation. Subsequent tests in 1965 and 1966 indicated that the items in the final instrument and the discriminatory formula separated the subjects into their appropriate occupational groups on the expressed motivational level.

A NUMBER of studies have attempted to isolate factors related to choice of teaching as a career. The variety of instruments for analysis and approaches considering interests, attitudes, preferences, aptitudes, chance factors, cultural pressures, and values as determinants have failed to provide a valid and reliable criterion for measurement and a basis for differentiation, selection, and prediction of teaching as a career.

Concerning the problem, Strong (12) asserted that the inventoried interests of men who are engaged in various occupations differ significantly from those of men in general, and he could not differentiate business executives and teachers from the men in general group. Super (13) pointed out that executives and teachers are not truly occupational groups and that "teacher" is too broad a category for psychological study. Studies concerning personality attributes and inventories were not helpful in the selection of teaching as a career because instrumentation fails in the assessment of personality factors such as needs, motivations, reaction tendencies, and adjustment (14). Furthermore, studies on specific patterns of traits or needs did not show clearly the differentiation of members of one occupation from those of another (15).

In the past, success in teaching has been emphasized as an important criterion for selection. Predictions of success have been unreliable when based on IQ scores of student teachers (9) or on correlations with standardized tests (8). Ghiselli (4) summarized data which make it clear that aptitude tests are better predictors of success in training than success on the job. As pointed out by Thorndike (16), beyond survival in an occupation, success is a meaningless concept, especially long-range success, which depends so heavily upon contingent factors that can never be predicted from what is known about the individual in advance.

The investigator questions the assumption of many behavioral scientists that individuals will not or cannot accurately and reliably estimate their own characteristics. Other investigators have found significant relationships between self-estimates and objective measures in intelligence (18), values (6), personality (2), and test scores (1, 10). In fact, Norton (7) reported that his data produced support for the validity of the ratings which were obtained by self-estimates.

The relative predictive value of inventoried interest and expressed preference has been studied by Wightwick (19), who concluded that measured interests are not as valid a predictor of vocational choice as are expressed preferences. Holland (5) appeared to bear this out when he reported that there is clear evidence that students' reports of their (vocational) choice are more predictive than any other evidence available. Super and Bachrach (14) seemed to be in agreement with these remarks. Stickney, Gaier, and Moore (11) used the technique of expressed motivations to assess motivations for counselor training. They reported the technique to be valid. It is interesting to note that their design almost parallels the design of Cohen's study (3). The new Thorndike Dimensions of Temperament (TDOT) test uses this approach in order to organize the respondent's acknowledged preferences and self-perceptions into a useful summary profile (17).

Problem and Procedure

This study was designed specifically to determine whether or not expressed motivations for curricular choice would discriminate between male college seniors enrolled in business administration and male college seniors majoring in secondary education and engaged in student teaching. This type of population eliminates variables stemming from mixed responses of males and females, and limits the variables operating from other than the secondary school level. Seniors, who are still identified as a group most likely to enter the profession, were chosen in order to avoid variables associated with attitudinal changes due to actually working on the job. This research does not attempt to identify or seek to use unconscious or non-expressed motivations, restricting itself to those readily stated by the subject and herein referred to as "expressed motivations." Based on these findings, the investigator attempted: (a) to develop a usable formula for the differentiation and selection of prospective teachers; (b) to establish a technique to survey other kinds of vocational choice; and (c) to discover a more effective instrument to employ in personnel and guidance.

Two groups of samples, randomly drawn from New York University (N.Y.U.) School of Education and School of Commerce and Queens College, New York City, were given the following topics to be completed in essay form: "Why did you choose to enroll in teacher education?" or "Why did you choose to enroll in business administration?" The students' statements in response to these questions were taken at face value and accepted without interpretation. At one stage in the investigation, these responses were kept in the respondents' words. Later, the responses took the form of marks on a checklist.

Although women classmates were given the essay question along with the men (to avoid reactions against being "singled out"), the data on females (F) were not used. The difficulty in obtaining data about males in the educational field required the extension of sampling to a period of 2 years. Complete anonymity of the response was promised in order to obtain accurate data. Then open-end preliminary interviews were held with these students to explore the motives expressed for making their curricular choice.

In order to determine whether the motivations expressed by these groups were significantly different, a list of all expressed motivation was compiled. As shown in Table 1, the original essays contributed 558 separate items. This was reduced to 289 items on the final questionnaire. The reduction was accomplished by removing duplications and by using a single statement that covered two or three similar statements. Two judges checked on duplications selected for elimination. Final selected items were weighted and combined into a questionnaire with a 5-point intensity scale in order to enable the subject to indicate how strong a motivation an item was in his choice. Each subject was asked to circle 0 to indicate no motivation; 1 to indicate slight motivation; 5 to indicate strong motivation; 2, 3, or 4 to indicate intermediate degrees of motivation.

To test items, this questionnaire was administered to a new sample of male business adminis-

Table 1.—Summary of Sample Sizes

Procedure	Reason	Number of Cases		Totals		Percent Responses Returned
		Business	Education	Cases	Items	
Essay[a] Distribution (1963-1964)	To amass items for questionnaire	326	178	504	558	81
Questionnaire[b] Distribution 1 (1965)	To test items	97	43	140	289	37
Questionnaire[b] Distribution II (1966)	To check reliability	55	25	80	289	30
	Totals	478	246	724		

[a] Approximate time required to complete essay - 15 minutes.

[b] Questionnaire required a minimum of 30 minutes to complete.

Table 2.—t-Ratio and Factor Analysis

Statistics Used	Purpose	Result
t ratio	To locate significant items.	From 289 original items, 125 were significant at .05 and an additional 30 were significant at .10.
Factor Analysis I	To locate clusters of items significant at .05 level as determined by t ratio.	12 factors were located. Because of unreliability of last 6 factors, 6 were retained. 122 items total.
Factor Analysis II	To locate additional clusters by adding items significant at .10 level as determined by t ratio.	Of the 30 items added, 8 were found to cluster significantly. 96 items total.
Factor Analysis III	To combine all items found to be significant in locating factors.	4 major factors were isolated. The 2 other factors were dropped because of too few items. 89 items retained.

tration seniors and male student teachers. Items that discriminated between administration seniors and student teachers were identified. From the items that discriminated, a scoring key was devised so that the questionnaire could be used as a measuring device.

Results and Discussion

Discriminant analysis was chosen as the proper statistical method of treating the data. The criterion variable was the curricular choice, and the expressed motivations were used as the prediction variables. Reasoning that not all 289 variables were significant, a t ratio was used to locate significant variables. According to Table 2, 125 items were found to be significant at the .05 level and thirty items at the .10 level.

In order to find out whether these significant items indicated any degree of cluster, factor analysis was applied as a "sorting" of the common characteristics inherent in the test. This enabled the investigator to isolate the common underlying "teacherness" and "businessness" within the groups. To be certain that all factors were located and that all items of significance were given an opportunity to cluster into a factor, the factor analysis was run three separate times. As shown in Table 2, Factor Analysis III located four major factors with cluster characteristics that were highly significant ($P < .001$). These four factors were composed of eighty-nine separate variables, but, since they clustered together as factors, the separate factors could be programmed into the discriminant equation.

Once placed into the discriminant-analysis equation, the weights of factors were calculated according to the number of factors programmed. These results were recalculated three times, using four, three, and two factors for consideration each time in obtaining new weights (see Table 3). Since the

Table 3.—Discriminant Analysis

Statistic Used	Purpose	Result
Discriminant Analysis I	To develop the proper equation weights using 4 factors.	1 (-.00041) 2 (+.00090) 3 (-.00061) 4 (-.00044)
Discriminant Analysis II	To develop the proper equation weights using 3 factors	1 (-.00042) 2 (+.00090) 3 (-.00075)
Discriminant Analysis III	To develop the proper equation weights using 2 factors	1 (-.00044) 2 (+.00085)

Table 4.—F Ratios

Statistic Used	Purpose	Result
F ratio I	To determine significance of Discriminate Analysis I. (This analysis used all 4 factors.)	F(4,135) = 63.366**
F ratio II	To determine significance of Discriminant Analysis II. (This analysis used 3 factors.)	F(3,136) = 84.014**
F ratio III	To determine significance of Discriminant Analysis III. (This analysis used 2 factors.)	F(2,137) = 124.215**

** $P < .001$.

F ratios showed an improvement with each successive analysis (see Table 4), Factors 1 and 2 were selected as the important factors for the equation of this study, using the factor weights developed from these final two factors (Table 3) as the coefficient weights for the final discriminant equation.

The above position is further strengthened by the following analyses. Table 5 shows the mean values for each factor. Means of both business students and education students were developed from their responses. The table shows large differences between these means for Factors 1 and 2. The differences between the means for Factors 3 and 4 are statistically significant, but small. It remains to be seen just what contribution Factors 3 and 4 may yet make.

Table 5.—Mean Values and Differences for Factors

Factor	Mean of Business Students	Mean of Education Students	Difference
1	73.59794	127.79070	-54.19276
2	84.19588	40.76744	43.42843
3	19.09278	19.39535	-.30257
4	4.96907	8.04651	-3.07744

Table 6.—Discriminant Analysis for Factors 1 and 2

Factor	Sample Size	Mean v	Variance v	SD
1	97	.03861	.00045	.02113
2	43	-.02226	.00043	.02070

Table 6 shows the v (discriminant analysis) for Factors 1 and 2. Note the small variances, suggesting little spread within the factor about the means. Table 7 indicates a list of items that successfully discriminated student teachers and college seniors enrolled in business administration, and presents a "profile" of individuals oriented toward teaching and business careers. Statistically, these items were referred to as Factors 1 and 2.

Based on the findings, the following equation was developed to quantify the separation between groups. The discriminant function coefficients arrived at were: —.00044 and .00085. This means that the discriminant equation may be written as follows:

$$v = \begin{array}{c}\text{factor 1 coefficient}\\ (-.00044)\end{array} \times \begin{array}{c}\text{(score for}\\ \text{factor 1)}\end{array} + \begin{array}{c}\text{factor 2 coefficient}\\ (.00085)\end{array} \times \begin{array}{c}\text{(score for}\\ \text{factor 2)}\end{array}$$

If the result (v) is .00817 (average of the Education Mean v and Business Mean v) or higher, the subject is a business student. If the result is .00817 or lower, the subject is an education student.

To test the items in the final instrument, the questionnaire was administered to a new sample of ninety-seven male business administration seniors and to forty-three male student teachers from N.Y.U. and Queens College. (See Tables 1 and 6.) The results are as follows:

a. Among the ninety-four Ss with v's of .00817 or higher, ninety-one were business students and were correctly identified from among a total of ninety-seven business students.

b. Among the forty-seven Ss with v's of .00817 or lower, forty were education students and were correctly identified from among a total of forty-three education students.

In the above test, six students were found to be placed in cross-groups, i.e., three students whose scores were more typical of business than teachers were taking student-teacher training, and three students whose scores were more typical of teachers were taking business courses. Although this represents only six cases out of 140 (approximately 4 percent), it does not necessarily indicate imperfection in the equation. A possible explanation may be sought in the inappropriate choice of occupation by these specific individuals. If such an assumption proves to be correct, then proper identification and counseling could be of value not only to this particular group of students but also to improve the results of the test. In the same way, counseling could also be of value to those sixty-three students whose scores fell in the

Table 7.—Item Profiles of Prospective Teachers and Prospective Businessmen

Item No.	Statement
	Factor 1 - "Profile" of Prospective Teachers
12.	Joy and psychological remuneration.
14.	I am concerned with social-moral events and conditions.
16.	It allows me to have idealism.
41.	I feel a deep satisfaction when I impart to others what I know.
42.	In this field, I feel more singly responsible for the products of my efforts.
54.	I could raise the standard of living of those less fortunate than me.
59.	It offers me a chance to travel abroad to learn how other people live.
62.	I can get closer to students than most others in this field.
67.	The world situation warrants the need for my chosen field.
71.	To help young people to mature intellectually is a great and noble task.
73.	I want to make a place where kids can grow.
78.	It is the most rewarding of all professions.
79.	Literature, language, critical judgment, and logical and emotional sensitivity can offset this race towards self-destruction.
81.	It offers me an opportunity for self-analysis.
94.	I enjoy children.
95.	I was always willing to share my knowledge.
124.	I believe that I am contributing something toward the improvement of situations in neighborhood, state and nation
130.	I am interested in the development and improvement of man.
133.	It allows people to express themselves in a beautiful way.
143.	I have always enjoyed reading and writing.
145.	After every lesson, the students respond and somehow seem a bit more enriched spiritually.
146.	It offers me more time for personal reflection.
154.	It stimulates my mind to relate outside interests and experiences to this field.
163.	It offers me an opportunity to work with and affect people.
165.	I have knowledge to give that would aid others in the maturation of their personalities.
172.	This field can offer more to a youngster in developing his needs and abilities than any other.
175.	There aren't enough people really interested in the needs of others.
185.	It offers a challenging and dynamic outlet for creativity.
187.	It offers me an opportunity for self-expression.
188.	It offers me an opportunity to convey my beliefs to as many people as possible.
189.	It would allow me to make kids assume their role in society.
190.	The experience of seeing how much prejudice existed among people.
192.	It would allow me to do something useful.
198.	I could play an active part in many processes of discovery.
202.	I am concerned for what seems to be an increasingly perilous situation.
205.	This field is the foundation upon which we must build.
206.	It is more beneficial in all walks of life.
209.	I believe that students going through their formative years need guidance and understanding.
210.	This field needs men who have a firm understanding of it.
223.	This field is a major part of a person's life.
226.	It would allow me to prepare others to face the reality of life.
227.	I could further the individual development of my children and fulfill their potential in society.
229.	It would allow me to become involved in the academic world.
235.	It would allow me to help others.
240.	The teacher must believe in his power to teach.
270.	I was happiest during my years in high school.
	Factor 2 - "Profile" of Prospective Businessmen
2.	It would help me to gain independence.
10.	It would enable me to get a better job.
25.	It is a field in which the monetary advancement is based upon ability.
37.	It presented the most advantages upon graduation.
47.	It is a wide open field.
64.	It is competitive and I like to compete.
68.	The financial advancement of the position appeals to me.
77.	I have the ability to adapt to various situations.
90.	It is a growing field with numerous opportunities.
91.	The field offers me numerous opportunities.
93.	It would allow me to make a lot of money and still be young enough to enjoy it.
96.	This field is the care of all corporations.
97.	It offers flexibility in a career.
100.	It is a field that offers a wide selection of areas I can successfully pursue.
107.	I feel more convinced that I will eventually end up in business.
111.	In this field I could advance myself more readily.
114.	A person with something on the ball, a "sharp" person, can make much money in this field.
118.	Money, recognition, and satisfaction are only to be gained by leading.
135.	I want to get an understanding of how our economic and business system operates.
147.	I enjoy fast-paced living.
168.	I have an interest in big business.
169.	It is a field which is important to the economy of the country.
171.	Regardless of the field you study, most people will enter the business world.
207.	It offers more money and bigger opportunities than do most other fields.
228.	It will serve as a basis for an eventual executive position.
236.	It is a field which requires analysis and I enjoy analyzing.
255.	The best possible chance to go to school and keep abreast of what's happening on Madison Avenue.
265.	Opportunities for advancement were limited without a degree.

middle ground between the means of the two groups (approximately 45 percent). At this point in the research, the computer equation seems to show that nearly one-half of the students enrolled in business administration or teacher education may not be sure of having made the proper choice for their future profession even up to the time of their graduation. With the instrument developed in this study, it may now be possible to identify and to assist students in their occupational choice.

The final equation was check-validated 1 year later with another sample of students. The questionnaire was administered to fifty-five male business administration seniors and twenty-five male student teachers from N.Y.U. and Queens College (see Table 1). The questionnaires from these new groups were scored, and the scores for the two groups were compared. The results are as follows:

> a. Among the fifty-two Ss with v's of .00817 or higher, all were business students and were correctly identified from among a total of fifty-five business students.
>
> b. Among the twenty-eight Ss with v's of .00817 or lower, twenty-five were education students and were correctly identified from among a total of twenty-five education students.

It is of particular interest to note that the validity study was able to identify correctly 97 percent of the sample. The extremely high prediction scores attest to the underlying reliability of this instrument and of the approach developed in this study. It is also important to note the underlying stability in the equation itself over a period of a year.

Conclusions

Based on the results of this research, it may be concluded that: (a) The expressed motivations for curricular choice did discriminate between seniors enrolled in business administration and teacher education; (b) The discriminatory equation developed in the study separated the Ss into their appropriate occupational groups, which defied Strong's efforts, and prompted Super to label them as "not true occupational groups."

If these conclusions are valid, the approach outlined in this investigation appears to supply a tool and technique of analysis with which to explore and to make expressed motivations more productive than expressed interests. The differentiation between student groups suggests that similar research on the expressed motivations of employed groups may likewise prove to be useful and instructive.

Even though this study has distilled some of the most important underlying motivations that distinguish and suggest that "teachers" have some motivations in common which are in some way different from those motivations which "businessmen" have in common, nevertheless a note of caution should be given as to the limitations of this study. The samples used represent limited geographical and cultural groups of one sex only. The instrument developed can be used in its present form as a preliminary screening device (for counselors and advisement) in a specific geographical area and population. The applicability of this study to wider segments of the collegiate population and geographical area must be subjected to further validation and reliability.

REFERENCES

1. Adamek, E. G., "The Effects of Testing and Test Interpretation on Selected Self-perceptions," unpublished doctoral dissertation, University of Illinois, Urbana, 1961.
2. Amatora, M., "Validity in Self-evaluation," *Educational and Psychological Measurement*, 49:119–126, 1956.
3. Cohen, Daniel, "Expressed Motivations in Curriculum Choice," unpublished doctoral dissertation, New York University, New York City, 1967.
4. Ghiselli, E. E., *Measurement of Occupational Aptitude*, University of California Press, Berkeley, 1955.
5. Holland, J. L., "Some Explorations of a Theory of Vocational Choice and Achievement. II. A 4-year Prediction Study," *Psychological Reports*, 4: (Monograph Supplement 12), 1963.
6. Nickels, J. B.; Renzagalia, G. A., "Some Additional Data on the Relationship Between Expressed and Measured Values," *Journal of Applied Psychology*, 42:99–104, 1955.
7. Norton, J. L., "A Study of the Estimates of Worker Trait Requirements," *Personnel and Guidance Journal*, 41:59–61, 1962.
8. Rolfe, J. F., "The Measurement of Teaching Ability, Study No. 2," *Journal of Experimental Education*, 14: 52–71, 1945.
9. Seagoe, M. V., "Prognostic Tests and Teaching Success," *Journal of Educational Research*, 38:685–690, 1945.
10. Singer, S. L.; Stefflre, B., "Analysis of Self-estimate in the Evaluation of Counseling," *Journal of Counseling Psychology*, 1:252–255, 1954.
11. Stickney, T. H.; Gaier, E. L.; Moore, G. D., "Motivational Factors of N.D.E.A. Institute Enrollees Reflected in Autobiographical Data," *Vocational Guidance Quarterly*, 14:44–50, 1965.
12. Strong, E. K., Jr., *Vocational Interests of Men and Women*, Stanford University Press, Palo Alto, California, 1943.
13. Super, D. E., *Appraising Vocational Fitness*, Harper and Brothers, New York, 1949.
14. Super, D. E.; Bachrach, P. B., *Scientific Career—Vocational Development Theory*, Bureau of Publications, Teachers College, Columbia University, New York, 1957.
15. Super, D. E.; Crites, J. O.; Hummel, R. C.; Moser, H. P.; Overstreet, P. L.; Warnath, C. F., *Vocational Development: A Framework for Research*, Bureau of Publications, Teachers College, Columbia University, New York, 1957.
16. Thorndike, R. L., "The Prediction of Vocational Success," *Vocational Guidance Quarterly*, 11:179–187, 1963.
17. Thorndike, R. L., "Thorndike Dimensions of Temperament Test," *Psychological Corporation Test Service Bulletin*, 56:1–11, 1966.
18. Webb, W. B., "Self-evaluation; Group Evaluations; and Objective Measures," *Journal of Consulting Psychology*, 19:210–212, 1955.
19. Wightwick, M. D., "Vocational Interest Patterns," *Teachers College Contribution to Education*, No. 900, 1945.

Forthcoming Articles

DeBlassie, Richard R.; New Mexico State University, Las Cruces
"A COMPARATIVE STUDY OF THE PERSONALITY STRUCTURES OF PERSISTENT AND PROSPECTIVE TEACHERS"

Dobbs, Ralph C.; University of Missouri, Columbia
"SELF-PERCEIVED SOLUTIONS TO ADULT PROBLEMS"

Lambert, Philip, Lee H. Hansen, Edgar F. Borgatta; The University of Wisconsin, Madison
"INTELLIGENT WORD ASSOCIATIONS IN HIGH SCHOOL STUDENTS"

Linton, T. E. and K. I. Altman; University of Wisconsin–Milwaukee
"OPERANT CONDITIONING IN THE CLASSROOM SETTING: A REVIEW OF THE RESEARCH"

Nelson, Deane D.; Moorhead State College, Moorhead, Minnesota
"A STUDY OF PERSONALITY ADJUSTMENT AMONG ADOLESCENT CHILDREN WITH WORKING AND NONWORKING MOTHERS"

Otto, Wayne, Thomas C. Barrett, Richard J. Smith, Kenneth L. Dulin, Dale D. Johnson; The University of Wisconsin, Madison
"SUMMARY AND REVIEW OF INVESTIGATIONS RELATING TO READING, JULY 1, 1969 TO JUNE 30, 1970"

Ross, Louis; University of Akron, Akron, Ohio
"FORECASTING THE ACADEMIC ACHIEVEMENT OF ENGINEERING FRESHMEN"

Strag, Gerald A. and Peraline P. Yeatts; University of Georgia, Athens
"FLEXIBILITY OF COGNITIVE STYLE AND ITS RELATIONSHIP TO ACADEMIC ACHIEVEMENT IN THE FOURTH AND SIXTH GRADES"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 5, January 1971)

Effects of Group Counseling on Creativity

S. JEFFREY GARFIELD
Educational Resources, Inc., Chicago

HELEN A. COHEN, ROBERT M. ROTH, HARRIS L. BERENBAUM
Illinois Institute of Technology

ABSTRACT

The purpose of the present study was to investigate the possible relationships between improvement in group counseling and changes in assessed creativity. It was hypothesized that creativity is positively related to mental health and that subjects who were judged improved after 30 hours of group counseling would also show improvement on variables typically associated with creativity. Thirty-eight subjects who participated in a group counseling experience over a 15-week time span were pre- and posttested on various measures of creativity, and pre- and post-rated on Rogers' elements of creativity, i.e., (1) Openness to experience; (2) Locus of evaluation; (3) Ability to toy with ideas. At the end of the experimental time the subjects were assigned to either an improved or unimproved category on the basis of therapists' ratings, outside judgments, and Q adjustment scores. The results indicate that improvement in group psychotherapy correlates significantly with gains in ratings on the more dynamic aspects (creative life style) of creativity. The correlations with the more cognitive measures of creativity were in the predicted direction but not significant. Further research was suggested.

SINCE 1952 "creativity" has become a very popular area of research. The plethora of publications in recent years attests to this. Presently, three major themes emerge from the wealth of material on this subject.

1. Creativity is viewed as a cognitive factor and is related to intelligence.
2. Creativity is viewed as a correlate of personality organization.
3. Creativity is viewed in its relationship to mental health.

This third approach to creativity intrigues personality theorists. Certainly a composite description of the creative individual, as reported in the empirical literature, presents him as emotionally healthy. Various studies have found the creative individual to be: independent, original, inner-directed, tolerant of uncertainty, intelligent, clever, imaginative, socially independent, and emotionally stable (1, 2, 3, 5). Few psychologists would argue with the proposition that this aggregate of traits represents an emotionally healthy individual.

Rogers (8) offers a theoretical rationale for these empirical findings. He states that creativity involves self-realization, a process requiring three elements. 1. *Extentionality,* or openness to experience, which permits the individual to tolerate ambiguity and relax psychological defensiveness; 2. An *internal locus of evaluation* which allows the creative person to form his own judgment of his products and personal worth; and 3. The *ability to toy with elements and concepts.* Rogers views these elements of creativity as important aspects of self actualization and in the Rogerian system a self-actualizing individual is emotionally healthy. The question then arises: What is the specific relationship of creativity to mental health? Is creativity synonomous with healthy, is it a function of emotional maturity, or is it a cognitive function which may or may not be related to personality structure. The purpose of this study was to investigate the relationship between creativity and mental health, specifically the relationship of increased mental health (as measured by a variety of clinical criteria) to increased creativity (as measured by a battery of ratings and cognitive tests).

To accomplish this an attempt was made to foster creativity by means of personality modification using group therapy.

Method

Subjects. Ss were forty-seven male undergraduate students at an urban university between 17 and 21 years of age.

Procedure. Thirty-eight Ss were given two group-centered therapy sessions a week over a 15-week period. Nine Ss served as a wait therapy control group. After therapy, all Ss were assigned to improved or unimproved groups. At the beginning and end of the 15-week period, all Ss were given a series of creativity tests and were rated on Rogers' elements of creativity, i. e., (1) Openness to experience; (2) Locus of evaluation; (3) Ability to toy with ideas.

Rating Procedures. Pre-therapy: Taped diagnostic interviews with each subject were rated by three clinical psychologists on Rogers elements of creativity using a 7-point scale. The median between raters (percentage agreement between raters was .70) was used to determine the Ss scores on each element.

Post-therapy: The same rating procedure was used for the elements of creativity. In addition each subject was rated improved or unimproved in mental health by his therapists and was also rated improved or unimproved in creativity in general on a 7-point scale.

Testing Procedure. The subjects were given pre- and post-therapy, Guilford's Uses for Things test, Barron–Welsh Art Scale and two Q sort instruments.

Guilford's Uses for Things: This test assesses "ideational fluency" and "semantic spontaneous flexibility." Ss were required to name as many different uses as possible for a given common object (e. g., brick). Each S was scored for both originality and productivity of uses.

The Barron-Welsh Art Scale: Ss were asked to rate (like or dislike) sixty-two black and white drawings. Barron (3) has summarized the personality correlates of this scale: Artistic preference is negatively correlated with rigidity, repression, social conformity, and impulse control. This scale has successfully discriminated creative from non-creative artists, architects, writers, etc.

The Q-Adjustment Score (7), based on the Q-sort was devised to provide an external criterion of adjustment level. Each was instructed to choose from among a set of cards those most descriptive of him, those least descriptive, and finally those about which he was indifferent.

Trained and practicing clinical psychologists determined which statements the well-adjusted person would say were most like or unlike him. The tally of the Ss distribution of responses, as compared to the "ideal" well-adjusted person, was called the "adjustment score."

The Self-Sort-Ideal-Sort Discrepancy Score (7) was used as an internal criterion of self-acceptance. Subjects given the same cards as above were instructed to select those most like themselves (self-sort). Next they were told to sort the cards describing the ideal person they wished to be (ideal-sort).

It is believed that the discrepancy between the placement of the given characteristic on the self scale and the ideal scale would yield an indication of self-esteem. The true-false self-ideal sort was found to correlate .90 with the type of Q-sort procedure described for the Q Adjustment Score (4).

Selection of Groups. After therapy, all Ss were assigned to "improved" or "unimproved" groups. For an S to be rated "improved," he must have been (1) rated improved by his therapist, (2) in the upper 50 percent of improvement scores on the self–ideal Q sort, and (3) in the upper 50 percent of improvement scores on the Q Adjustment Test. All Ss rated as unimproved by their therapist and who failed to meet at least one of the Q sort improvement measures were in the "unimproved" group.

Experimental Group I: Nineteen Ss rated improved in therapy.

Experimental Group II: Nineteen Ss rated unimproved.

A Wait Therapy Control Group consisted of nine Ss who were randomly deleted from therapy due to the lack of an opening in the program.

Results

As can be seen from Table 1, Mann Whitney U Tests showed that mean improvement scores of the psychologically improved group (Experimental Group I) were significantly higher at the .05 level of confidence on ratings of (1) openness to experience and (2) locus of evaluation that the mean improvement scores of the psychologically unimproved group (Experimental Group II) or the Wait Therapy Control Group.

There were not significant differences between the psychologically unimproved group (Experimental Group II) and the Wait Therapy Control Group on these improvement ratings.

Table 1.—A Comparison of Improvement on the Creativity Variables of the Improved, Unimproved, and Wait Therapy Control Groups

Creativity Improvement Variables	Improvement vs. Unimprovement	Improvement vs. Wait Control	Unimprovement vs. Wait Control
Openness to Experience	2.88**	2.58**	.28
Locus of Evaluation	2.59**	2.28*	.28
Ability to Toy with Elements	1.44	1.23	.02
Therapists' Rating of Creativity	3.00**	***	***
Original Uses Test	1.59	.27	1.15
Different Uses Test	1.51	1.08	.19
Barron Welsh	.80	.73	-.15

* .05 level of confidence
** .01 level of confidence
*** Rating not available on Wait Therapy Controls

Therapists' ratings of improvement in creativity were significantly higher at the .01 level of confidence for the psychologically improved group (Experimental Group I) than for the psychologically unimproved group (Experimental Group II). Because the Wait Therapy Control Group did not receive therapy, therapists' ratings of improvement in creativity were impossible to obtain for that group.

The only improvement rating variable that the psychologically improved group did not show significant gains on, as compared to the other two groups, was the Ability To Toy With Elements and Concepts. The facts that (1) the raters experienced great difficulty in trying to separate open toying from defensive toying, and (2) this variable had the lowest inter-judge reliability, may account for the lack of significant differences on this variable. The ratings were felt to measure the more dynamic aspects of creativity. Considering the fact that the psychologically improved group showed significant gains on three of the four rating variables seems to support the hypothesis that improvement in group therapy is positively correlated with improvement in certain dynamic aspects of creativity.

As can also be seen from Table 1, there were trends in the expected direction on improvement scores of the psychologically improved group (Experimental Group I) versus the psychologically unimproved group (Experimental Group II) and the Wait Therapy Control Group on virtually all of the tests used to assess cognitive aspects of creativity. However, these results were not significant at the .05 level of confidence.

Summary

This study suggests that increases in mental health as a result of group therapy may also be expected to result in increases in several dynamic aspects of creativity. Trends of increases in cognitive aspects of creativity were also noted for the group that improved in mental health via group therapy, but these increases were not statistically significant as compared to the other two groups. Further research seems warranted on attempting to increase cognitive aspects of creativity as well as on the relationship between cognitive and dynamic aspects of creativity.

REFERENCES

1. Barron, F., "Some Personality Correlates of Independence of Judgment," *Journal of Personality*, 21:287-297, 1953.
2. Barron, F., "Originality in Relation to Personality and Intellect," *Journal of Personality*, 25:730-742, 1957.
3. Barron, F., "The Psychology of Creativity," in Newcomb, T. M. (ed.), *New Directions in Psychology II*. Holt, Rinehart and Winston, New York, 1965.
4. Garfield, S. J.; Cohen, Helen, A., "A Simplified Method of Using the Self-ideal Q Sort," unpublished research, Illinois Institute of Technology, 1967.
5. MacKinnon, D. W., "Genus Architectus Creator Varietas Americanus," *American Institute of Architects Journal, September*, 1960, pp. 31-35.
6. MacKinnon, D. W., "Fostering Creativity in Students of Engineering," *Journal of Engineering Education*, 52:129-142, 1961.
7. Rogers, C. R.; Dymond, Rosalind F. (eds.), *Psychotherapy and Personality Change: Co-ordinated Studies in the Client-centered Approach*, University of Chicago Press, Chicago, 1954.
8. Rogers, C. R., *On Becoming a Person*, Houghton Mifflin Co., Boston, 1961.

ders

Titles from DERS

RESEARCH AND DEVELOPMENT TOWARD THE IMPROVEMENT OF EDUCATION
Edited by Herbert J. Klausmeier and George T. O'Hearn—$5.75

In this book twenty eminent scholars, researchers, and research administrators present their latest thinking about the processes and conditions of human learning, the processes and programs of instruction, and the application of research and development strategies to educational problems.

CANADA GOOSE MANAGEMENT
Current Continental Problems and Programs
Edited by Ruth L. Hine and Clay Schoenfeld—$5.50

The Canada Goose *(Branta canadensis)* is the center of widespread interest on the part of waterfowl managers, hunters, resource administrators, and the wild-life-enjoying public. This book assembles comprehensive current data on Canada geese from all parts of North America. It contributes significant information on the present problems, new research, and future prospects of Canada geese on a continental basis.

Order from: DEMBAR EDUCATIONAL RESEARCH SERVICES, INC.
Box 1148, Madison, Wisconsin 53701

FIELD NEWS

Willard E. North
Central Missouri State College

A 4-YEAR WORK STUDY program for incoming freshman majors in elementary education has been initiated at Fairleigh Dickinson University and will be under the direction of Julius O. Luck, Professor of Education and Director of Environmental Education.

The program, which is on a voluntary basis, gives the freshman the opportunity to participate in the elementary classroom situation in his first semester and at the same time permits him to decide whether or not his interests are geared for the teaching profession.

The program begins with a teacher's aide position offering the student 2 years of experience, under the direction of a master teacher. The teaching aide experience is supplemented by a 2-hour on-campus seminar-workshop under the supervision of a member of the faculty. The seminar-workshops will be dedicated to specific areas and the full sequence will be designed to provide each participant with an introduction to the background and organization of public education in the United States, and to an understanding of the relationship between human growth and development and the teaching–learning process, of the community society and the culture in which children and youth live and learn, of the organization in operation of schools instructional program. It will also provide a knowledge of instructional materials and their uses in the teaching process, and opportunities and procedures for student's self-evaluation in accordance with recognized and validated performance criteria.

At the completion of the 2-year program, the student can enter an internship. During this period, he will be observed by school supervisors on a regular basis for evaluation purposes. In the senior year, the student can participate in a full time teacher-in-training program where he or she will normally teach a full schedule of classes under the supervision of a master teacher and the college faculty.

* * * * * *

AS REPORTED IN the October issue of ECS Bulletin, published by the Education Commission of the States (ECS), legislators in at least twenty-two of the fifty states are ready with proposed measures designed to curb violence on the campus. These findings are the results of a nationwide survey completed recently by the ECS.

During 1969 and 1970, bills and joint and concurrent resolutions designed to prevent campus disorders were introduced and adopted in thirty-two states. Stiff fines and jail sentences were prescribed in much of the legislation with some, as in the case of Oklahoma, ranging as high as $15,000 or 10 years imprisonment for such offenses as rioting and burning or destroying property.

Legislation "relating to promotion of or participation in riots" since 1969, and proposed for consideration by legislatures in 1971, includes measures to enable university authorities to name security officers and invest them with full police powers to prepare rules and regulations for maintenance of order, and to ban from their campuses persons who do not measure up to these rules.

Still other anti-campus-riot legislation revokes scholarships and financial aid to students who take part in violence, and sets up controls on campus speakers, and simpler restrictions. In some states anti-riot legislation reflects specific demonstration or anti-war techniques. For example, a new Oklahoma statute prohibits anti-war demonstrators from publishing or displaying on placards on public property the names of members of the armed services killed in action without first obtaining the written consent from the surviving spouse or next of kin.

In California, which had the unhappy distinction of having more campus unrest in 1969 than any other state, 120 bills and resolutions have been introduced in the legislature in 2 years. Most of them were ultimately combined into two substantive bills which enable junior college, college, and university authorities to take a number of disciplinary actions, including dismissals, withdrawal of financial aid, and similar actions against any student, faculty, or other staff member convicted of a crime arising out of a campus disorder.

Pennsylvania legislators introduced more than a dozen student and faculty unrest bills in their 1969 and 1970 sessions and while only one of them became law, the ECS survey indicates that it may well have the greatest effect throughout the country. The far-reaching Pennsylvania bill provides that each approved institution of higher education, wherever it may be, in the United States or abroad, shall furnish to the state's Higher Education Assistance Agency "the name and address of any student who is a resident of the Commonwealth of Pennsylvania who is expelled, dismissed, or denied enrollment for reasons set forth in this section . . . or of whom the institution of higher learning has knowledge that he has been convicted of offenses . . ." so his financial aid can be stopped. At the time of the survey, only three Pennsylvania and thirty out-of-state schools on an approved list of 2,600 have indicated that they would not sign an agreement to report their Pennsylvania students.

Bills which were introduced in the Connecticut legislature but failed to pass included one that would have required the immediate expulsion of any student convicted of participation in a campus disorder and another that would have made such students ineligible for state financial aid. Four factors were identified as reasons why the bills did not pass: (1) the public and private institutions formed a united and responsible opposition to the proposed legislation; (2) public confidence in higher education institutions was strong enough to permit them to maintain their academic autonomy; (3) there were few incidents of campus unrest at the state institutions; and (4) the legislature's education committee was respected, had provided outstanding leadership, and refused to support the proposed bills.

* * * * * *

THE ERIC CLEARING HOUSE on Educational Management (formerly the Eric Clearing House on Educational Administration) has announced the availability of a special series of selected bibliographies and brief analyses of literature on eleven topics related to school organization and administration. These publications were prepared during the summer of 1970 for a new research and development program in the United States Office of Education. Publications now available are: (1) *Educational and Social Demands on the Schools;* (2) *Alternative Organizational Forms;* (3) *New Sets of Jobs for School Personnel;* (4) *Management Information Systems;* (5) *Models for Planning;* (6) *Models for Rational Decision Making;* (7) *Procedures for Managing Innovations;* (8) *Linking Schools to State Education Departments;* (9) *Linking Schools and State Education Departments to Research and Development Agencies;* (10) *New Programs for Training School Administrators;* (11) *Introduction Plans (Marketing Plans).*

For a free copy of any or all of these publications write to Editor's Office, ERIC Clearing House on Educational Management, University of Oregon, Eugene, Oregon 97403.

* * * * * *

AUTHORS OF REPORTS of research and other projects concerned primarily with tests, measurements, and evaluation are urged to send duplicate copies to ERIC Clearing House, Educational Testing Service, Princeton, New Jersey 08540. The most timely and significant reports will be abstracted and indexed for publication in *Research in Education,* a guide to research information issued monthly by the Educational Resources Information Center (ERIC) of the U. S. Office of Education, Washington, D. C. If the author can abstract his own report in 200 words or less, so much the better. The report itself will be sent to ERIC for reproduction in microfiche or hard copy on request. If measurement is not the primary concern, some reports may be sent to other ERIC clearing houses. Those of local or limited value will be retained at ETS, discarded, or returned to the author at his request.

* * * * * *

THE 7-PART *SEL Pathways to Better Schools Series* is now being distributed. The results of 3 years of work with rural isolated school systems by the Southeastern Education Laboratory (SEL), the Pathway Series was designed to serve as a guide in planning and preparation of project proposals.

The second generation "Pathways" includes six illustrative Title III proposals and a Comprehensive Planning Guide which may be used by a local school system to assess its current program and resources as a part of the planning process. Numbers 2 through 7 use actual Title III proposals, written and submitted by local school systems, as illustrative cases. Each case includes a commentary by the laboratory which outlines Title III requirements for Alabama, Florida, Georgia, South Carolina, and Mississippi. A list of suggested references, an outline of what the ideal statement should encompass, and a critique of the illustrative case accompany each section of the proposal.

The seven segments of the Pathways series are: (1) Comprehensive Planning Guide; (2) Organization for Instruction Program; (3) In-Service Training Program; (4) Reading Program; (5) Dropout Reduction Program; (6) Preschool Training Program; and (7) Communications Skills Program.

Copies of each segment are available free of charge. A limited supply has been produced and requests will be filled upon written request until the supply is exhausted. The Laboratory asks that readers limit their requests for those documents which will be of greatest use. Requests should be directed to Southeastern Education Laboratory, 3450 International Boulevard, Atlanta, Georgia 30354.

* * * * * *

THE VARIATIONS IN SPEECH (dialects) used by some Americans indicate only that the speaker grew up in a certain area of the country or within a certain social class. However, to the listener (if he grew up somewhere else) the dialect may indicate that the speaker is ignorant, uneducated, unfit to hold a job, or even hostile. These are some of the findings of a study carried out by Charles Billiard, currently Director of English Education at Georgia State University in Atlanta. Billiard's study (carried out at Purdue) indicates that the inability of some black Americans and Latin Americans to speak the dialect of the middle class is a real barrier in getting an education and a better job.

Subjects from Fort Wayne, Indiana, were selected for the study because of the influx of southern Negroes since World War II and the more recent addition of many Latin Americans to the community. The investigator tape-recorded interviews held with fifty residents of Fort Wayne and Allen County. He then analyzed the tapes to pick out significant words and usages. Interviewed Ss were classified by ethnic group and by social class which was determined by education, occupation, and income.

The speech of the upper-middle class, both blacks and whites, was taken as the standard dialect. Two other dialects were identified. Some of the more obvious dialect differences included pronunciation and stress variations. For example, a Fort Wayne resident of the upper-middle class pronounces the "r" in "for" but the black newly arrived from the South usually pronounces the word as "foe." Southern blacks also place the emphasis in some words differently than do upper-middle class whites and blacks, thus the "poLICE" of the upper-middle class becomes the "POlice" of the Southern blacks.

Latin Americans also tend to have some differences of emphasis. The upper-middle class speaks of the "RAILroad." The Latin calls it "railROAD." The Latins also have a limited range of vowel sounds at their command. The Anglo-American "this" comes out "thees," for example.

After the investigator had identified the variations in the dialects, he made new tapes in which important variant words such as "for" and "police" were read as part of sentences by different speakers who gave them their variant dialect pronunciations. Then Billiard played the tapes for audiences representing black-American, Anglo-American, and Latin American ethnic groups in the upper-middle class, middle class, and lower class social groups. Other audiences included personnel managers and supervisors from businesses and industry and elementary and high school teachers.

Billiard found that upper-middle class Anglo-Americans and teachers were most sensitive to variations from the standard dialect. Upper-middle class black-Americans were most tolerant of variances.

In one phase of the study, an upper-middle class black was incorrectly identified as Anglo-American by more than nine of the ten listeners who heard him speak in a standard Fort Wayne dialect. When the same man imitated southern Negro dialect, the listeners were equally emphatic in their identification of the speaker as a Negro.

Billiard suggests that the inability of minority group members to speak the standard dialect of the community presents new challenges to English and elementary school teachers. He also feels that teacher education should give more emphasis to the study of social dialects and the problems of motivating students to learn a second dialect. He observes that teaching the second dialect may be much like teaching a foreign language. A student first must learn to hear the differences in the standard speeach and his "native" tongue. He then must learn to produce the standard sounds and practice until he becomes fluent.

Among Those Present

DANIEL COHEN is an Assistant Professor of Educational Psychology at Pace College, Director of the New York Testing and Guidance Center, and holds a consultative appointment as Director of Visual Psychology at the Institute of Visual Physiology and Optics. He is biographied in Who's Who in The East and the American Psychological Association's Directory of Personnel, and is listed in the National Vocational Guidance Association's Professional Member Directory, and in Community Leaders of America. He holds three degrees (BA, MA, PhD), all from New York University. He has taught 11 years on the secondary school level, and has been on the full-time staff at Pace College since 1967.

JOHN S. DACEY is Director of the Educational Psychology Division as well as the Honors Program, of the School of Education, Boston College, Chestnut Hill, Massachusetts. He received his BA from the State University of New York at Binghamton; his MEd and PhD were awarded by Cornell University. **GEORGE F. MADAUS** is an Associate Professor of Education in the School of Education, Boston College, Chestnut Hill, Massachusetts. He received his BA in Sociology from the College of the Holy Cross; his MEd was awarded by the Massachusetts State College at Worchester, and his EdD was awarded by Boston College. He was a post-doctoral fellow at the University of Chicago in 1965–1966.

GLENN DAHLEM earned his doctorate in Counseling and Guidance under the late Wray Strowig at The University of Wisconsin, Madison. Somewhat of a generalist in his approach to school counseling, he has worked as a secondary, elementary, and K-12 counselor, as well as a college counselor-educator. He is presently an elementary counselor with the Pattonville, Missouri, schools and is writing a book on pupil personnel services. His primary interest area for research is counselor role and image, including both self-image, and concepts held by various publics of school counselors.

S. JEFFREY GARFIELD has a PhD in Clinical Psychology from Illinois Institute of Technology (IIT) and is currently serving as President of Educational Resources, Inc. **HELEN A. COHEN** is Director of the counseling-community psychology program at IIT. She received the PhD from the University of Chicago. **ROBERT M. ROTH** is an ABEPP psychologist who is Associate Professor of Psychology at IIT. He recently authored the books *Underachieving Students and Guidance* and *The Psychology of Vocational Development*. **HARRIS L. BERENBAUM** is Assistant Professor of Psychology and Director of the counseling center at IIT. He received the MS and PhD from IIT.

CHARLES B. KEENEN is presently Dean of Student Services at Germanna Community College, Fredericksburg, Virginia. At the time of this study he was Assistant Dean of Men at the University of Florida. He received his AB from Dartmouth College; MEd from Springfield College, and EdD from Boston University. He has held positions in admissions and counseling and is particularly interested in the selection and retention of students.

WILLIAM E. LOADMAN was a research consultant in the Office of Research Consultation in the College of Education at Michigan State University. He is completing his doctoral dissertation and will receive his degree in Measurement, Evaluation, and Research Design in Educational Psychology. He has taught research and statistics at Michigan State University. His research interests include learning, pattern analysis of interest inventories, and evaluation of ongoing projects. His professional experience includes teaching, evaluation of federal projects, and full-time research consultant for the Michigan Department of Education with respect to the statewide assessment of basic skills in the fourth and seventh grades during the 1969–1970 school year.

JAMES B. MACDONALD received his doctorate from The University of Wisconsin, Madison. He has taught in the elementary schools and at the University of Texas, Austin, New York University, and The University of Wisconsin, Madison. He is currently Professor and Chairman, Department of Curriculum and Instruction, The University of Wisconsin—Milwaukee.

JOSEPH PILAND is President of the Lincoln Trail Junior College at Robinson, Illinois. He holds an EdD in Educational Administration from Illinois State University. He has been both an elementary principal and dean of general studies. **ELMER LEMKE** is an Associate Professor of Psychology at Illinois State University. He holds the PhD from The University of Wisconsin where his training was in psychometric methods.

RALPH L. SHAW holds a BA from Georgia State College and an MA from Atlanta University. He has done graduate work at both Emory University and the University of Georgia. He is now acting director of the Office of Testing and Evaluation of Emory University. **NORMAN P. UHL** received his PhD from the University of Maryland. He was director of the Office of Testing and Evaluation and Professor of Education and Psychology at Emory University. Later, he was a research psychologist with the Educational Testing Service in Durham, North Carolina. Presently, he is a Professor at North Carolina Central University in Durham.

C. LELAND SMITH is Associate Professor of Education at the University of Kentucky. He holds the AB in Music, MA in History, MA in Education, and the EdD in Educational Administration. He has had experience in the public schools as a teacher and principal. He formerly taught at Transylvania University, Lexington, Kentucky. Most recently he has served as director of several NDEA History Institutes. He is presently teaching in the department of curriculum and has special interests in the psychosocial adjustment of students.

RICHARD G. WEIGEL received his PhD in Counseling Psychology at the University of Missouri, Columbia, and is currently Assistant Professor of Psychology at Colorado State University. He has contributed a number of articles to the professional literature in counseling psychology, and has recently co-authored the extended monograph, *The College and University Counseling Center*. **FRED McKINNEY** is Professor of Psychology, University of Missouri. He received his graduate education at the University of Chicago and is the author of four books, numerous journal articles, a video course in general psychology, and an assortment of video cases, interviews, and demonstrations for the psychology of personality. He was formerly president of Division II (Teaching) of the American Psychological Association, and is a former chairman of the Conference of State Psychological Associations, a Division of the American Psychological Association.

CAROL WIESNER holds a BA from Sarah Lawrence College, an MEd from the University of Hartford, and a PhD in Educational Psychology from the University of Connecticut. She has worked as a teacher, curriculum coordinator, and psychological consultant in the public schools. Presently, she is on the University of Hartford faculty serving on a project designed to develop educational services for the children of migrant farm workers in Connecticut.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor of Curriculum and Instruction, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Associate Professor of Special Education, Yeshiva University, New York.

GEORGE S. JURATA, Coordinator of Reading, Jefferson County Schools, Colorado.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

S. JAY SAMUELS, Associate Professor of Educational Psychology, University of Minnesota, Minneapolis.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Editors Emeriti

HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan 48075.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

CARLETON W. WASHBURNE, Director Emeritus of Teacher Education and the Division of Graduate Studies, Brooklyn College, 2248 Kent

Diary No. 289
Date 16.6.71
File No Lib
Bureau Edn'l. Psy. Research

JeR

50th Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 FEBRUARY 1971 NUMBER 6

ARTICLES

Summary and Review of Investigations Relating to Reading, July 1, 1969 to June 30, 1970—*Wayne Otto, Thomas C. Barrett, Richard J. Smith, Kenneth L. Dulin, Dale D. Johnson, Robert Barganz, Susan Hawk, Juanita Peters* 242

Intelligent Word Associations in High School Students—*Philip Lambert, Lee H. Hansen, Edgar F. Borgatta* 269

The Schools' Contribution to the Cumulating Deficit—*Sunnyuh Shin Chang, James Raths* 272

Operant Conditioning in the Classroom Setting: A Review of the Research—*K. I. Altman, T. E. Linton* 277

DEPARTMENTS

Field News—*Willard E. North* 287

Forthcoming Articles 268

Book Reviews—*Robert E. Clasen* 271

(The Journal of Educational Research is indexed and/or abstracted in Abstr.S.W., CIJE, CSPA, Current Contents, Ed. Adm. Abst., Educ. Ind., Lang. and Lang. Behav. Abstr., M.H.R.Ind., PsychoAbstr., Soc. of Ed. Abst.)

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, of Editorial Services

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 6, February 1971)

Summary and Review of Investigations Relating to Reading, July 1, 1969 to June 30, 1970

WAYNE OTTO, THOMAS C. BARRETT, RICHARD J. SMITH, KENNETH L. DULIN, and DALE D. JOHNSON, The University of Wisconsin—Madison
with the assistance of ROBERT BARGANZ, SUSAN HAWK, and JUANITA PETERS

THE CURRENT Summary and Review is, as has been the policy in recent years, limited to investigations reported in journals. One of the main purposes of the review is to reflect recent trends in the published research related to reading. Consequently, the review is somewhat selective and the internal organization in terms of sub-topics may change from year to year. The investigations continue, however, to be grouped by area: sociology, psychology, physiology, and teaching of reading; a separate bibliography is given for each of the four main areas. As usual, the placement of a certain investigation in a given area may be rather arbitrary, so the reader is advised to peruse the entire review. While there is no systematic attempt to present a critical analysis of each investigation summarized, the authors of the review have felt free to make comments that clearly reflect personal perceptions. Our hope is that these comments can in all cases be distinguished from the objective summary statements.

For reviews of the reading literature designed to serve somewhat different and varied purposes, see those appearing with some regularity in such sources as *The Reading Research Quarterly, The Review of Educational Research, Elementary English, The Reading Teacher, The Journal of Reading,* the revisions of *The Encyclopedia of Educational Research,* and the proceedings and special publications of such organizations as the International Reading Association, the National Reading Conference, and the National Council of Teachers of English. Comprehensive summaries of research reports can, of course, also be obtained from ERIC-CRIER.

I. THE SOCIOLOGY OF READING

A. Environmental Influences Upon Reading

1. *Family Patterns and Reading Achievement.* To further extend a recent line of investigation relating ordinal birth-position to intellectual and academic potential, Oberlander and others (14) examined interactions between birth order, family size, and scholastic aptitudes and achievement within a sizable sample of suburban eighth-graders. Grade point average (GPA), SRA High School Placement Test subscores, and a derived scholastic aptitude score (IQ) were used as the bases for the analyses. IQ scores were significantly related to both birth order and birth order × family size in interaction, with these relationships decreasing as family size was lowered. Overall, it was concluded that the birth order × family size interaction was more meaningfully related to scholastic aptitude and achievement than were either birth order alone or family size alone. In relation to reading achievement, the birth order × family size interaction was significant ($p < .05$) and birth order alone closely approached significance ($p < .10$).

In a more broadly-based study, Goldstein and others (6) examined relationships between reading and several other school performance measures and a large range of family structure variables: socioeconomic status of family, employment of the mother, discord in the home, ambitiousness for the child, birth order, number of siblings, "density" (nearness in age) of siblings, and a combination index of maternal working-status and intra-home discord. Multivariate analysis indicated that mother's ambitiousness, father's ambitiousness, birth order, and ordinal position within the family were generally major predictors of school performance. The variable of major importance among the dependent measures was reading, with father's occupation, number of siblings, mother's ambitiousness, and father's ambitiousness in turn univariately correlated with it. The results of this study, though varying in terms of the types of analyses made, definitely indicate that any comprehensive attempt to predict school success should take into account family variables as well as measures designed to assess intellectual potential alone.

Another attempt at relating home and family factors to reading achievement was a study by Miller (12). On the basis of the data collected during in-depth home interviews, the typical daily schedules of kindergarten-aged children from middle, upper-lower, and lower-lower class homes were compared, rated in terms of specified criteria, and correlated with Metropolitan Readiness Test (MRT), Gilmore Oral Reading, and Stanford Achievement reading scores at the first-grade level. "Desirable" activities were such things as interaction with other members of the family, opportunities to develop experiential background, and activities developing auditory and visual discriminatory abilities. Significant group differences in daily schedules were found among the children of each of the three social classes, though within each group there were individual children participating in typical home activities not representative of their social class group as a whole. In general, children's daily home activities were not significantly correlated with either readiness or first-grade achievement scores in reading. Only in the lower-lower class group were significant correlations found, and there only with the readiness and oral reading scores, not with those of silent reading achievement.

Finally, a study by Cramer and Dorsey (5) tested the widely-held opinion that a family pattern of frequent moving has an adverse effect upon a child's reading achievement. High and low mobility groups were determined on the basis of family and school data for 366 sixth-graders from Air Force personnel families and compared in terms of age, IQ, reading comprehension, and reading vocabulary. Though age and IQ were significant contributors to reading achievement, no significant differences were found between high and low mobility groups. Actually, "high movers" scored consistently but not significantly *higher* in reading vocabulary and comprehension than did their less mobile classmates.

2. *Social and Cultural Disadvantage and Reading.* Several studies dealt with attempts to understand and/or

ameliorate the relationships between cultural deprivation and reading. Karnes and others (8), for example, reported upon the comparative success of four parallel programs of intervention into the intellectual and linguistic development of disadvantaged 4-year-olds: (1) a "traditional" nursery school program emphasizing personal, social, motor, and language development through informal learning situations, music and art activities, and indoor play with toys and dolls; (2) a "community-integrated" program characterized by the mixing of disadvantaged children with advantaged peers used primarily as language models; (3) a Montessori program built around Montessori society approved materials, "practical life" demonstrations of grooming and mealtime behaviors, and supervised field trips; and (4) a highly structured experimental program making use of manipulative and multi-sensory materials to elicit verbal behavior thought to be conducive to language development. Prime among the experimental program's activities were book-related and pre-reading behaviors, such as relating pictures to context, developing left-to-right progression, and the association of printed symbols with meaning. Each group was approximately two-thirds Caucasian to one-third Negro and approximately one-half males and one-half females. Evaluation of the four 7-to-8 month programs was done by pre- and posttesting with three instruments: Form L-M of the 1960 Stanford-Binet, the 1961 experimental edition of the Illinois Test of Psycholinguistic Abilities, and the Peabody Picture Vocabulary Test. The experimental and traditional groups moved in a positive direction on all instruments, with the experimental group's gains generally larger and achieved with greater consistency. In contrast, the community-integrated and Montessori groups showed smaller and less consistent gains. The experimenters concluded that the tying of verbalizations to motor-sensory performance was the crucial characteristic for a successful program of improving school readiness abilities. The relative absence of language activities in the highly structured Montessori program was deemed its major weakness, and the community-integrated program was felt to have failed because the disadvantaged children—a one-to-two minority in their group—failed to interact reciprocally in verbal situations with their more advantaged peers.

Johnson and Jacobson (7), too, reported on an ameliorative program dealing with disadvantaged learners: kindergarten-level, first grade, and second grade Negro children classified as ill-prepared for reading on the basis of reading readiness and/or reading achievement test scores. A 5-week summer program was designed to promote conceptualization and verbal growth predicated upon five major themes: (1) the building of experiential background through field trips, (2) the building of critical listening skills through story telling and discussion, (3) the creating of meaningful oral language experiences through the use of pupil-dictated experience charts, (4) the sharpening of auditory and visual discrimination skills through word games and poetry, and (5) the broadening of conceptual background through role-playing activities coming out of occupational and recreational story materials. The results of the study indicated that the program was successful at all three grade-levels in developing verbal conceptualization, met with some success at the kindergarten and first-grade levels in developing general verbal skills as measured by readiness and listening inventories, and was helpful at the second-grade level in developing reading readiness. The most striking gains were on conceptualization ability, as measured by the Peabody Picture Vocabulary Test, with mean increases in IQ points of 17.5, 19.3, and 8.3 at the three grade-levels.

In contrast to these two programs, Cawley and others (4) reported follow-up research on Head Start Program and non-Head Start Program children which indicated little or no gains for those having completed the program. They concluded that, without planned follow-through, programs like Head Start are simply unable to completely "catch up" children suffering from intense deficits in conceptual and experiential background. They suggest that perhaps intervention programs must begin much earlier than age 4 or 5 in a disadvantaged child's life, perhaps even as early as age 18 months.

Somewhat related to the Karnes and others hypothesis (i.e., that "the tying of verbalizations to motor-sensory performance is crucial in improving school readiness abilities") were a series of experiments reported in detail by Scott (21, 22). Based upon the results of testing done with an experimental Seriation Test developed by the investigator and his associates, the Piaget-Inhelder concepts of *seriation* (ordering by size and/or pattern) and *classification* (placing into named categories) were linked to the more common labels of *perception* and *language* and investigated as they relate to reading readiness and achievement, general intellectual ability, race, and social class. At the kindergarten level, Seriation Test scores were highly stable and closely correlated with MRT variables. Of particular interest was the finding that seriation scores were more closely associated with reading readiness than with number readiness, thus supporting the Karnes and others position that perceptual and language capabilities are highly interdependent and interactive, and that language enrichment, unaccompanied by perceptual experiences, fails to meet the young child's full requirements for sound intellectual growth and development. As had been hypothesized, Negro children in this study scored lower than whites on the Seriation Test, but a predicted social class effect did not occur, perhaps, the experimenter notes, because of possibly inappropriate pairings of the school populations involved. As a predictor of reading success at the third-grade level, seriation scores functioned with statistically significant power for all groups except Negro females, though predictive power was consistently lower for either-sex Negroes than for whites. In terms of prediction, a final comparison of interest was that of teacher's rankings at the kindergarten level, Seriation Test score rankings, and total MRT rankings. For most subgroups, all three functioned about equally well, but for Negro females teacher ranking was the very best predictor, while for Negro males it was the poorest.

Another race-related study was done by Saunders (20). Though fundamentally an attempt to assess the cross-racial validity of two projective techniques, the House-Tree-Person and Draw-A-Person tests, it dealt secondarily with reading and is therefore of interest here. From a total of 240 Negro and white first graders of the same socioeconomic level, who had been carefully evaluated in terms of IQ, reading achievement, and emotional adjustment, three pools of forty pupils each were set up: ten white poor readers, ten white good readers, ten Negro poor readers, and ten Negro good readers in pool one; ten white well-adjusted pupils, ten white poorly-adjusted pupils, ten Negro well-adjusted pupils, and ten Negro poorly-adjusted pupils in pool two; and ten bright whites, ten dull whites, ten bright Negroes, and ten dull Negroes in pool three. Three judges, each of whom held the PhD in psychology, then made four sets of judgments on the basis of the children's drawings, judging each child in terms of the three stated categories and also in terms of race. Judgments made concerning the three major categories of classification were all highly accurate, above the .02 level of significance, while judgments of race were only 50.8 percent accurate, almost exactly what would be predicted by chance. The author concluded from this that the two projective techniques are indeed valid cross-racially and not biased across the two ethnic groups as long as socioeconomic levels are comparable.

A particularly interesting study, difficult to review but perhaps the most engaging of all those examined by the reviewer, was a biracial study reported by Lewis and others (10). The term "reported" is used here advisedly, since that is exactly what the investigators did: they *reported* the data and let the conclusions speak for themselves. Two reading groups (adequate and inadequate) within two racial groups (Caucasian and Negro) were studied in terms of twenty-six variables: verbal, performance, and full-scale Weschler's Intelligence Scale for Children (WISC) IQ's; scores on each of the twelve WISC subtests; a socioeconomic level (SEL) index; a motor-development index; a symbol-coding test; a visual memory test; a free-response writing productivity test; three family situation variables; age; and an Iowa Test of Basic Skills reading criterion measure. As would be expected, the two reading groups differed significantly on almost all (17 to 19) of the intellectual and academic variables, but the racial groups did *not*. The only WISC variables, in fact, where the two racial groups differed were in the performance subtests, the portion of the WISC testing, the authors point out, which was done first, while test anxiety was still high and testee-examiner rapport low. On family situation variables, the racial groups were clearly different, with Negroes of lower socioeconomic status (SES), larger families, and lower families, and lower father's education. In terms of other variables, Negroes were younger and more productive of language, and whites were higher on visual memory and symbol-coding. For this reviewer, at least, the overall conclusion reached was that social differences, not academic and/or intellectual ones, were the only telling biracial ones.

Two interrelated studies by Bruininks (2, 3) dealt with the auditory perceptual and visual perceptual skills and reading performances of disadvantaged Negro boys. In the first study (2), a battery of six auditory perception tests and six visual perception tests were administered to 105 primary-school level Ss, and the Ss' individual performances on these measures were then related to their reading achievement and verbal intelligence scores. Significant correlation ratios and/or correlation coefficients with reading achievement were obtained on only six of these tests: (a) the Wepman Auditory Discrimination Test, (b) Digit Span, (c) the Roswell-Chall Auditory Blending Test, (d) Auditory Attention Span for Related Syllables, (e) the Children's Embedded-Figures Test, and (f) the Visual Automatic Test. Generally, the auditory discrimination tests correlated more highly with reading than did the visual perception measures.

In the second study (3) Bruininks attempted to fit the teaching of word recognition to the dominant (i.e., visual or auditory) perceptual strengths of each S, but with limited success. The Ss learned to recognize words equally well under teaching methods matched to either their strengths or their weaknesses. Overall, presentation stressing visual factors was more successful, rather surprising in terms of the first study's results.

A final study by Becker (1) dealt with the use of language experience methods in a Job Corps Center. Though his N of only four and lack of hard pretest data hardly justify the inclusion of his study in a collection of empirical research, his report does stand out as a fine stage-by-stage description of a promising remedial methodology.

3. *Socioeconomic Levels and Reading.* A particularly comprehensive factor-analysis study by Uhl and Nurss (24) investigated what each of the tests in a battery of fifty-three perceptual, cognitive, linguistic, and achievement measures was assessing in two socioeconomic level groups: upper-middle and low. The Ss were 211 second-grade boys and girls from the Atlanta Public Schools, approximately one-third Negroes and two-thirds whites from each SEL. Correlations among the fifty-three variables plus SEL, race, and sex were initially calculated, but only six were found to be highly (.30 or above) related to race, and none to sex; therefore, factor analyses were made only in terms of SEL. Several well-defined factors were found common to both groups, as well as an additional factor for each separate group. Visual perceptual speed and accuracy, visual discrimination of words and letters, visual motor coordination, immediate memory span, academic achievement in the language arts, and vocabulary factors were found to be operating in both groups, with a separate auditory processing factor present in the low SEL group and a separate verbal mediation factor present in the upper-middle SEL group. The authors suggested that these differences may have been influenced by differences between a nonstandard and a standard American-English dialect predominant in each of the two groups. Two of the measures included in the factor analysis were reading achievement tests: Reading—Isolated Words and Reading—Paragraph Meaning. Both of these factors loaded on the academic achievement factor for both SEL groups. For the upper-middle SEL children these variables also loaded on the vocabulary factor, while for the lower SEL group the loading was related to oral language factors such as rhyming and listening comprehension. Thus, the authors concluded, the two groups of children appeared to be using different skills in approaching the reading tasks, one more cognitive and the other more oral. This suggested to the authors that, when teaching SEL children to read, a teacher should give maximum attention to these oral language factors and be aware of possible differences between the children's nonstandard American English dialect and her own.

Also dealing with vocabulary factors and SEL was a study by Labercane and Armstrong (9), who measured the recognition and recall vocabularies of high and low SES Canadian children and then evaluated each group's choices of response qualitatively. In comparing the two groups for both recognition and recall vocabulary, the high SES children gave a significantly higher number of abstract responses, and the low SES group a significantly higher number of error responses. For the total sample, however, SES levels did not significantly predict recognition vocabulary scores. Thus, the authors concluded that the low SES child's ability to recognize words within a given context is higher than is his ability to call up word meanings solely from memory, and that programs of remediation for such children ought probably to begin at this position of relative strength.

4. *Urbanity–Non-urbanity and Reading.* Two studies seem to fall into this somewhat questionable category. Libby (11) compared the word-recognition skills, reading comprehension, and total reading scores of inner city, outer city, and "transfer" (recently moved outward) children of matched intelligence. Differences between the three groups were all significant beyond the .01 level except between transfer and inner city children in comprehension, with the outer city group consistently superior to the transfer group, and the transfer group consistently superior to the inner city group. The author hypothesized that "transfer" families perhaps gain strength when able to leave the ghetto, or the decreased distraction of the outer city is somehow ameliorative.

The second study falling into this category dealt with the problems of children from a long-neglected, in the opinion of this reviewer, social and linguistic community: the isolated, though not necessarily economically deprived, semi-rural family. Contrasting these children with the inner city blacks or Appalachian whites who speak a truly nonstandard dialect, Skinner (23) analyzed language samples from children living in several selected rural Maine villages. These language samples, oral responses to various cartoons and pictures, showed more of a paucity of language flexibility rather than a truly nonstandard dialect. Relating this to a continuum of elaborate to restrictive language behaviors, Skinner characterized the speech of these children as simply quantitatively and qualitatively less mature than that of similar-age middle class urban and suburban children, but still within the limits of standard American English. Living in a society which does not place high priority on language, in other words, these children have simply not developed this facility to an appropriate level. As a teaching procedure to meet this problem, Skinner suggested language-experience methods.

5. *Delinquency and Reading.* Only one study appeared to fall clearly into this category. Reported by Rice (18),

it dealt with a broad-range approach to reading remediation within an institutional setting and was labeled by the author as "*Educo-therapy*," an "*eclectic, gestalt approach utilizing the disciplines of both psychology and education in an intensive treatment program*." Behavioral modification techniques were used both (a) to institute remedial education procedures, and (b) to stimulate improvement of self-concept and social integration. The ten Ss, all girls between the ages of 11 and 15, were characterized as follows: high in oral-pleasure needs, unable to postpone need-gratification, little impulsivity-control, low self-concept, lacking in humor concepts, rarely having had a good relationship with an adult, rarely having experienced success either at home or in school, and reading extremely poorly or not at all. No visual or auditory impairments were in evidence. Phase one of the treatment program focused primarily upon personal social and emotional growth, with a contingency system of oral and social reinforcement employed. Praise, affection, and individual attention were provided in addition to such oral reinforcers as chewing gum, candy, and food. Phase two added a full-day remediation program in reading and language skills development. Tokens were awarded for progress on a ratio basis, with the tokens then able to be used to "purchase" such things as candy, gum, food, and cosmetics. During phase three, the reinforcement schedule was revised, with greater and greater academic achievement and/or longer and longer periods of appropriate behavior necessary for the awarding of the tokens. Near the end of the program, tokens could be exchanged for off-campus excursions and visits home. At the end of the 3-month program, reading gains of from 2 to 13 months on the Gates Reading Survey had been attained, and IQ scores for individuals had risen from 1 to 17 points. Thus, it was concluded by the authors, the program was seen as a success both in terms of humanistic values and institutional economics.

B. Characteristics of Reading Materials

1. *American Newspaper Readability Levels.* After randomly selecting fifty metropolitan and fifty non-metropolitan daily newspapers and analyzing their formats and layout, Razik (17) applied the Dale-Chall formula to six sample articles from the front page of each. The results indicated that a significant difference in readability level existed between the metropolitan and non-metropolitan papers. The metropolitan papers were written at the ninth-tenth grade level and the non-metropolitan papers were written at the eleventh-twelfth grade level. Also, international news items were written at a higher level than were articles describing local events, crimes, and tragedies. The authors concluded that, since the median level of education in the United States is only 10.6 years, a good many important newspaper articles are regularly being written above the reading comprehension level of at least 50 percent of the adult reading population.

C. Reading Habits and Preferences

1. *Adult Communicative Activities.* Essentially replicating a pioneer 1930 study, Samovar and others (19) surveyed four hundred randomly selected adults in the San Diego metropolitan area to determine the relative time they spent daily in various communication activities. The returns were separated into various occupational groups and analyses done both in terms of communicating time generally and time spent in specific communicative activities. Retired people, students, and engineers logged the highest reading time: 21, 20, and 20 percent of waking time, respectively. When communication activities for the total group were rank-ordered in percentages of waking hours devoted to each, the following hierarchy emerged: conversation, 24.1 percent; speaking, 15.2 percent; listening, 13.1 percent; televiewing, 12.2 percent; and reading 11.1 percent. Without adjustment for ties, 50 percent of the occupational groups logged reading as their lowest communication activity; only engineers ranked it as highest.

2. *Media Use by Foreign Students.* To investigate the types of media used most heavily by foreign students attending educational institutions in the United States, Mowlana and McLaughlin (13) sampled fifty-nine foreign students at the University of Tennessee. The overall ranking found was (1) radio and television, (2) U. S. newspapers and magazines, (3) foreign publications, and (4) personal sources. Students in the humanities relied more on printed media than did those in engineering. The most universally read magazine (noted by 28% of the sample) was *Reader's Digest*.

3. *Professional Groups and Reading.* To determine the perceived usefulness of mass-media medical information to medical doctors, O'Keefe (15) questionnaired two groups: medical faculty at Duke University and the University of North Carolina, and every third physician listed in six North Carolina cities. When asked how often they received information about new developments within their own specialties from the mass media, only 30 percent of the doctors said this happened as often as once a month. Non-teaching doctors tended to receive this information more frequently than did faculty physicians. About the same percentage (11%) of both groups reported they never received information this way. Newspapers, with magazines a close second, were cited as the source from which they most often received useful medical information. While 71 percent considered the information received to be of sufficient reliability, only 24 percent said this information was of any importance to them. Non-faculty people in general found the information to be more important. There was a significant difference between the opinions of faculty and non-faculty doctors on the importance of making stories about medical developments available to the public, with faculty people considering it more important. For the total group, only about two-thirds thought it was even fairly important. From these comments, it would appear that the role of the mass media is relaying useful medical opinion to the profession and public, at least in the view of the medical profession itself, is minimal.

Patterson and others (16) examined the readership of science articles in daily newspapers. A collection of science stories, roughly one-third dealing with biological science, one-third with physical science, and one-third with social science, was assembled and given to a group of equally represented scientists, educated non-scientists, and communication people to rank on a scale of —3 to +3. The data thus collected were then factorially analyzed. Three groups emerged, roughly coinciding with the three groups represented in the sample. Group A had relatively high educational background and interest in science, but preferred their science stories "palatable" and easy to take. Group B had less education, but worked daily in some way related to science. This group was very content-conscious, and preferred biology articles over space ones. Group C was made up completely of communications people, and their preferences were consistently for "people-oriented" social science articles.

REFERENCES ON THE SOCIOLOGY OF READING

1. **Becker, John T.**, "Language Experience Approach in a Job Corps Reading Lab," *Journal of Reading*, 13: (no. 4)281–284, 319–321, January 1970, I.A.2.
2. **Bruininks, Robert H.**, "Auditory and Visual Perceptual Skills Related to the Reading Performance of Disadvantaged Boys," *Perceptual and Motor Skills*, 29: 179–186, 1969, I.A.2.
3. **Bruininks, Robert H.**, "Teaching Word Recognition to Disadvantaged Boys," *Journal of Learning Disabilities*, 3: (no. 1)30–37, January 1970, I.A.2.

4. **Cawley, John F.; Barrow, Will H.; Goodstein, Henry A.**, "Performance of Head Start and Non-Head Start Participants at First Grade," *The Journal of Negro Education*, 39:(no. 2)124–131, Spring 1970, I.A.2.
5. **Cramer, Ward; Dorsey, Suzanne**, "Are Movers Losers?" *The Elementary School Journal*, 70:(no. 7) 387–390, April 1970, I.A.1.
6. **Goldstein, Kenneth M.; Cary, Gene, L.; Chorost, Sherwood B.; Dalack, John D.**, "Family Patterns and the School Performance of Emotionally Disturbed Boys," *Journal of Learning Disabilities*, 3:(no. 1)10–15, January 1970, I.A.1.
7. **Johnson, Joseph Carlton; Jacobson, Milton D.**, "Current Trends in Negro Education and Shorter Papers," *The Journal of Negro Education*, 39:(no. 2)171–176, Spring 1970, I.A.2.
8. **Karnes, Merle B.; Teska, James A.; Hodgins, Audrey S.**, "The Effects of Four Programs of Classroom Intervention on the Intellectual and Language Development of 4-year-old Disadvantaged Children," *American Journal of Orthopsychiatry*, 40:(no. 1) 58–76, January 1970, I.A.2.
9. **Labercane, G. D.; Armstrong, R. D.**, "Socioeconomic Status and the Meaning Vocabularies of Children," *The Alberta Journal of Educational Research*, 15: (no. 4)225–233, December 1969, I.A.3.
10. **Lewis, Franklin D.; Bell, D. Bruce; Anderson, Robert P.**, "Reading Retardation: A Biracial Comparison," *Journal of Reading*, 13:(no. 6)433–478, March 1970, I.A.2.
11. **Libby, Mildred Post**, "Differences in Reading Performance of Inner and Outer City Children Matched for Intelligence," *Education*, 90:(no. 2)145–147, November–December 1969, I.A.4.
12. **Miller, Wilma H.**, "An Examination of Children's Daily Schedules in Three Social Classes and Their Relation to First-Grade Reading Achievement," *California Journal of Educational Research*, 21:(no. 3)100–110, May 1970, I.A.1.
13. **Mowlana, Hamid; McLaughlin, Gerald W.**, "Some Variables Interacting With Media Exposure Among Foreign Students," *Sociology and Social Research*, 53: (no. 4)511–521, July 1969, I.C.2.
14. **Oberlander, Mark; Jenkin, Noel; Houlihan, Kevin; Jackson, John**, "Family Size and Birth Order as Determinants of Scholastic Aptitude and Achievement in a Sample of Eighth Graders," *Journal of Consulting and Clinical Psychology*, 34:(no. 1)19–21, 1970, I.A.1.
15. **O'Keefe, M. Timothy**, "The Mass Media as Sources of Medical Information for Doctors," *Journalism Quarterly*, 47:(no. 1)90–100, Spring 1970, I.C.3.
16. **Patterson, Joye; Booth, Laurel; Smith, Russell**, "Who Reads About Science?" *Journalism Quarterly* 46: (no. 3)599–602, Autumn 1969, I.C.3.
17. **Razik, Taher A.**, "A Study of American Newspaper Readability," *The Journal of Communication*, 19: 317–324, December 1969, I.B.1.
18. **Rice, Ruth Dianne**, "Educo-Therapy: A New Approach to Delinquent Behavior," *Journal of Learning Disabilities*, 3:(no. 1)16–23, January 1970, I.A.5.
19. **Samovar, Larry A.; Brooks, Robert D.; Porter, Richard E.**, "A Survey of Adult Communication Activities," *The Journal of Communication*, 19:301–307, December 1969, I.C.1.
20. **Saunders, Mauderie Hancock; Teska, Percy T.**, "An Analysis of Cultural Differences on Certain Projective Techniques," *The Journal of Negro Education*, 39:(no. 2)109–115, Spring 1970, I.A.2.
21. **Scott, Ralph**, "Social Class, Race, Seriating, and Reading Readiness: A Study of Their Relationship at the Kindergarten Level," *The Journal of Genetic Psychology*, 115:87–96, 1969, I.A.2.
22. **Scott, Ralph**, "Perceptual Skills, General Intellectual Ability, Race, and Later Reading Achievement," *The Reading Teacher*, 23:(no. 7)660–668, April 1970, I.A.2.
23. **Skinner, Vincent P.**, "Implications for the Teaching of Reading to Non-Suburban/Urban Children," *Journal of the Reading Specialist*, 9:(no. 2)46–57, December 1969, I.A.4.
24. **Uhl, Norman P.; Nurss, Joanne R.**, "Socioeconomic Level Styles in Solving Reading-Related Tasks," *Reading Research Quarterly*, 5:(no. 3)452–484, Spring 1970, I.A.3.

II. THE PSYCHOLOGY OF READING

This year marks a substantial increase in the number of studies of language and perceptual factors included in the review and summary. Separate subsections are devoted to each area.

A. Psychological Correlates of Reading Achievement

1. *Intelligence.* Huelsman (42) reported an analysis of the WISC subtest patterns of 101 under-achieving and fifty-six achieving fourth graders from a non-clinical setting. In addition to reporting the results of his analysis, he reviewed twenty-three previous studies of WISC subtest patterns in the diagnosis of reading disability. The report would be worthwhile for the review alone; but the new data add substantially to what is known, or believed, about WISC subtest pattern analysis.

The prior fact, pointed out by Huelsman, is that a relatively well established pattern exists for poor readers: low information, arithmetic, and coding scores. He suggested that the value of such a pattern would be apparent in assisting diagnosticians in identifying cases of reading disability, in classifying them, and in teaching them efficiently. His review, however, yielded no direct support for the use of the subtest patterns with individuals. The main finding of the Huelsman study was an emphatic lack of support for the applicability of the low information, arithmetic, and coding sub-scores with individuals. Not one of the 101 poor readers studied scored low on all three subtests, and only six were low on two of the subtests. The author's conclusion, on the basis of this and related findings, was that ". . . research should be directed toward defining the possible significance of differences in WISC (subtest) scores rather than toward pattern identification which seems relatively useless." We agree with the "relatively useless" part, and wonder if reading diagnosticians could not be doing much more fruitful things than WISC sub-score analyses.

Henderson, Butler, and Goffeney (40) studied the effectiveness of the WISC and the Bender–Gestalt Test in predicting achievement in reading and arithmetic. The Ss, selected at random from among prenatal examinations at a county general hospital, were 120 white and eighty-three nonwhite children within 3 months of their seventh birthday at time of testing. Both the WISC and Bender scores were significantly related to the reading and arithmetic scores, but the Bender contributed little when WISC scores were partialled out. Craik and Masani (17) examined age and intelligence differences in young (age range = 19–33) and old (age range = 53–84) Ss' coding and retrieval of word lists. The conclusion was that the main effect of age is on retrieval, while coding is affected by verbal intelligence.

2. *Personality Factors.* A study in the continuing line of research on cognitive style was reported by Serafica and Sigel (77). The styles of categorization of twenty-four boys, age 8 to 10 years, with reading disablility were compared to the styles of twenty-four boys with no reading problems. The Sigel Cognitive Styles Test, which involves justifying the selection of two out of three categories that "go together" for thirty-five "sets" of three diverse categories, was administered to all Ss. The authors' hypothesis that the poor readers would give

fewer part-whole responses than the good readers was not confirmed, nor was the hypothesis that the poor readers would use more descriptive whole responses. The hypothesis that the poor readers would use more relational-contextual responses was confirmed. The authors concluded that the poor readers were not deficient in analytic ability, but that they differed from the good readers in conceptual and integrative abilities. The authors speculated that the poor readers' lacks might be caused by "neurological impairment or a social environment which does not promote conceptual development." Perceptual difficulties were dismissed as a cause with the observation that the responses of a perceptually impared subgroup did not differ from those of a perceptually adequate group.

Whitehill and Jipson (94) reported support for their hypothesis that introverts and extraverts would respond differentially to traditional versus instrumental reading programs. Extraverts responded more positively to an experimental speed improvement program, that featured highly structured attention-focusing conditions, than to a more traditional program.

3. *Perceptual Factors.* Lingren (51) matched twenty disabled readers with twenty "normal" readers on the basis of sex, IQ (90 to 110), and chronological age (8 to 14) and compared their performance on the Bender Visual Gestalt Test, the Wepman Auditory Discrimination Test, and a visual-motor matching task. The normal readers were superior to the disabled readers only on the Wepman test. The author suggested that lack of auditory discrimination may be a major contributor to the development of reading problems. While this is by no means a new finding, the study represents an attempt to control intelligence in demonstrating a difference in good and poor readers' performance on certain tasks. Golden and Steiner (33) also reported data which suggest that poor readers are lacking primarily in auditory rather than visual factors. Their Ss, ten pairs of good and poor readers matched on MA, IQ, and GA, responded to several subtests from the revised Illinois Test of Psycholinguistic Abilities and the Monroe Visualization test. All Ss were second graders.

Reed (66) put an interesting twist to his study of the relationship between copying tasks and reading achievement. He noted that a common approach is to start with disabled readers, often clinic referrals, and then to demonstrate deficiencies in their drawing and/or copying abilities. He reasoned that a causal relationship could be shown only by starting with children who have such deficiencies before they learn to read. Accordingly, the hypothesis tested was ". . . that groups of children who differ in copying ability would not differ in reading ability." Twenty-eight children whose copies of a Greek cross showed "failure to effect the proper spatial orientation" and twenty-eight children who produced satisfactory drawings served as Ss in the study. All Ss were first graders. Differences in reading achievement scores for the two groups were analyzed by an analysis of covariance to control for differences in WISC verbal and performance IQ's. There were significant differences between both the unadjusted mean reading scores ($p < .01$) and the adjusted mean scores ($p < .05$); but the difference in actual grade equivalents was so small that the author questioned the practical significance of the difference. The adjusted mean grade equivalent score of the good-Cross group was 2.5, as compared to 2.4 for the poor-Cross group. Correlation coefficients among IQ, copying, and reading were also significant; but again, when adjustments were made for IQ, the coefficient was so low (.26) that it was considered to have little practical importance. The author concluded that while the ability to copy a Greek cross amounts to one measure of intelligence, inability to copy is not a convincing explanatory concept for severe reading disability. A follow-up of Reed's Ss in another year or 2 would, however, be worthwhile, for the possibility of a delayed or cumulative effect still remains.

Parker (63) examined the musical perception of thirty-five poor readers and thirteen adequate readers at the junior high school level. His rationale for the study was stated mainly in terms of the demonstrated difference in auditory perception between good and poor readers. The main conclusion was that "musical perception in the usual sense of pitch discrimination and memory for tunes and rhythms is not a correlate of reading ability." There was some evidence, though, that lack of adequate rhythmic perception could hold up progress in reading.

4. *Multiple Factors.* Multiple correlates were examined in the two studies that follow.

Michal-Smith, Morgenstern, and Karp (56) did a detailed assessment of four siblings who were experiencing severe reading problems. Their stated threefold purpose was to (a) show the diversity of traits in reading disability, (b) determine whether reading disability is a contributing factor in emotional and social adjustment of the disabled reader and his family, and (c) explore the possibility of a congenital or developmental syndrome. The authors concluded that the four siblings did indeed have traits that fit their (the authors') definition of dyslexia, that heredity was the underlying causal factor, and that emotional problems compounded the total problem of reading disability. The theme of the report seems to be: the disabled reader's lot is not a happy one. This will probably come as no great surprise to most professionals in reading.

Rice (67) summarized the findings of a multidisciplinary study of children with learning disabilities. Multiple deficits were shown. More important, in terms of value of the article, it was suggested that a diagnostic clinic situated away from the educational setting has serious communication problems. The need for studies directed toward the development of teaching strategies designed to meet the problems identified was stressed.

5. *Preschool Factors.* The focus in several studies was upon selected antecedents of reading achievement.

O'Kelley (62) argued that because a child's comprehension vocabulary is critical to success in learning to read, an objective test of "listening vocabulary" ought to be of use in assessing readiness for reading instruction. His choice was the English Picture Vocabulary Test I (EPVT), a collection of forty items where the task is to "find the picture to go with word," which is very like such familiar tests as the Full Range Picture Vocabulary Test, the Quick Test, and the Peabody Picture Vocabulary Test. The main data reported show a high correlation ($r = .70, .71$) between the EPVT and the Stanford-Binet L-M scores of 209 children age 6.9 to 8.0, who had been identified as possible slow learners. The author concluded that the EPVT could be useful in screening for early detection of children in need of special help in their early school experience. Additional data in support of the existing norms for the EPVT are also reported.

The Screening Test of Academic Readiness (STAR) was evaluated by Magoon and Cox (53). STAR—which is composed of fifty items that cover picture-vocabulary, letters, copying, picture description, human figure drawing, relationships, and numbers—is a group test designed to discriminate between preschoolers who should and should not be accepted for early admission to school. The test was administered to a national sample of about 4,000 first graders and kindergarteners, and the data were subjected to careful analyses. The thrust of the report is succinctly stated by the authors in the last two sentences:

> *The STAR appeared to be measuring only lettering and numeration skills with any great degree of factorial validity, and this circumstance would probably not qualify the instrument as a readiness measure in developmental terms. Finally, since the STAR scores were found to be highly related*

to IQ scores, it is likely that there would be little difference between the screening potential of the STAR and an IQ measure.

The authors apparently did not have "STARs" in their eyes.

Ward (89) wondered if the Coding subtest of the WISC would be as good or better than the Murphy-Durrell Reading Readiness Test for predicting first grade reading achievement. Scores from both tests were obtained from five hundred kindergarten pupils and from the Stanford Achievement Test when 278 of the same pupils were completing first grade. The data showed that the Murphy-Durrell was a better predictor of all Stanford subscores than the Coding test. For word reading, for example, the coefficients were .28 and .45 respectively.

Hirst (41) questioned the validity of the assumption that school entrance age is a valid predictor variable for academic success. She reported a study designed to "identify the variables kindergarten children bring with them to school that may predict reading achievement through second grade." Children (N = 188) were followed through second grade, where data from achievement tests as well as a variety of predictive measures were available. Age did not turn out to be a significant predictor of first or second grade achievement. The author stated that this finding is in accord with research results reported in the last 5 years. She concluded that legislative action on entrance age does nothing more than eliminate the need for professional decisions.

Lyle (52) reported a study of selected antenatal, perinatal, and developmental variables as they relate to later reading performance. Subjects were fifty-four good and fifty-four poor readers, age 6 to 12, in Sydney. The most clear-cut relationship with reading disability was shown by three speech related variables: speech delay at 6 months, speech delay at 2 years, and early speech defects.

Wells (91) attempted to determine whether the preschool play activities of fourteen good and fourteen poor readers in third grade were related to their achievement in reading. The preschool data were collected from parents in a personal, structured interview. Differences favoring the poor readers were found in (a) amount of time spent outdoors in winter, (b) amount of time spent with mother in indoor activities, and (c) amount of time spent with father in outdoor activities. The suggestion was that it is not the kind of play activities but whether or not they are adult centered that appears to be critical. The author was careful to disclaim any demonstration of cause-effect, but she speculated that perhaps adult-centered activities created a dependency on adult interaction. Note should be made, too, of the fact that the author's rationale for the study gave reason to expect differences in preschool play activities. Wells' study is the kind that invites follow-up: the rationale is sensible, the method is creative, the reporting is credible, and the results are somewhat unexpected. We shall probably see more in this line of research.

6. *Prior Performance.* The studies reviewed here have in common their focus on the relationship of certain aspects of performance and subsequent gains in reading or reading related performance.

King, Dellande, and Walter (49) reported evidence to support their hypothesis that college students with beginning reading rates between 200 and 250 words per minute would make greater gains in GPA after completing a reading improvement course than faster or slower readers. They reasoned that the slower readers have complex problems and that the faster readers benefit relatively little from further improvement. Williams (96) attempted "to measure how much was comprehended by college freshmen before reading a passage and how much meaning was taken from the printed page after reading." Unfortunately, the data reported show only that different Ss bring different amounts of meaning to the printed page and that in most cases reading relevant material increases comprehension. The author's intent—to work toward predicting how much will be taken from a printed page when how much is brought to it is known—was not realized in any significant way. Thompson and Whitehill (86) reported support for their hypothesis that more flexible readers accrue greater speed gains from a reading improvement course than less flexible readers. They concluded that where flexibility is absent or limited, instruction designed to increase speed flexibility is a prerequisite to large and useful speed gains from speed improvement programs. Grob (34) demonstrated the excessive study time demands that prevail for secondary students with limited reading rates. He suggested differentiated assignments to reflect the reading rates of students. Roberts, King, and Kropp (68) used a vocabulary learning task to examine derivations from Ferguson's theory of human abilities. In general, there was support for the notion that different instructional methods would show different patterns of correlations with abilities over stages of practice and differential transfer.

B. Language Factors and Reading

1. *Oral Language.* Bougere (7) searched for ways of specifying those factors of oral language which are related to beginning reading achievement and for evidence that "selected language measures could add significantly to the predictive value of a standardized readiness test." A stratified random sample of sixty children from six suburban Chicago schools was tested on the Metropolitan Readiness Test (MRT) and on selected language measures based upon interviews and filmed illustrated stimuli. The Gray Oral and Stanford Primary 1 reading tests were used for criterion data. Major results showed that the MRT predicts reading success more accurately than any of the language measures, but the predictive validity of the MRT for comprehension and word recognition could be significally increased by the addition of three language factors: length of T-unit and range, and diversity of vocabulary.

Tatham's study (85) was concerned with determining second and fourth grade pupils' comprehension of written material containing frequent oral language patterns compared to their comprehension of material containing infrequent oral language patterns. Frequency of oral patterns was obtained solely from Strickland's findings. Two tests were constructed, each using thirty-seven sentences and 3-picture multiple-choice distractors. The Ss were randomly assigned to read either Test A (frequent oral language patterns: noun-verb agreement, negative-affirmative, etc.) or Test B (infrequent oral language patterns: manner, place, etc.) first; but took the other test the following day. Chi-square and t-test statistical procedures were employed. The principal finding was that significantly more second and fourth graders obtained higher scores on Test A than Test B ($p < .0001$). Tatham urges the use of children's patterns of oral language in written materials to increase comprehension.

Ostensibly, to determine the relationships between grade level, reading achievement, and oral language performance, Williams and others (95) estimated the grade levels of sixty pupils' oral language. Twenty Ss at each of three grades, 1, 3, and 5, responded to three stimuli pictures from the Chandler series in three ways. The oral responses were taped, transcribed, and assessed using the Botel Readability Formula. The Median Test was used to compute language scores for boys and girls. Oral language median scores were: 2.8 in grade 1, 3.00 in grade 3, and 3.05 in grade 5. The investigators concluded that girls make no gains in level of conversation during the first 5 years of school and boys gain only between first and third grade. The study's stated purpose is misleading, as no clear reference is made to oral language and reading level relationships. One could also question the adequacy of nine responses as the criterion for determining oral language level.

Using a single child's oral reading of a story from a basal reader, Burke and Goodman (10) very effectively illustrate the use of Goodman's Taxonomy of Cues and Miscues in Reading as an analytical tool for studying oral reading miscues. The S had 179 miscues, or 8.7 per 100 running words of text, which were analyzed in a variety of ways including: corrections, response function, expected response, type, level, graphemic, dialect, phonemic, syntactic acceptability and change, semantic acceptability and change, and intonation acceptability. The authors conclude that the taxonomy of reading miscues offers a framework against which to develop a very complete profile of an individual child's reading.

2. *Syntax.* Weisberg (90) reports two investigations designed to study sentence structure through intra-sentence word association. In Experiment I, thirty-five university undergraduates memorized three sentences with certain words underlined. Words from each sentence were then presented individually as word association stimuli, and students were asked to respond with one of the other words from the sentence. Experiment II was a refinement of Experiment I, in which the deep structure of alternate sentences was kept constant, while surface structure was allowed to vary. For example, the deep structure of the following sentences remains the same: "Slow children eat cold bread." "Slow children eat, as you know, cold bread." "Cold bread is eaten by slow children." Eighty university undergraduates comprised the sample. Both experiments were concerned with the general hypothesis that the most frequent response to each stimulus word would be more closely related to it in the deep structure of the sentence than in physical proximity in the surface sentence. The results from the two experiments suggest that the deep structure of sentences was the most important factor in word association, which suggests that as soon as a reader has dealt with a sentence, he is no longer responding to the surface structure.

3. *Semantics.* Svedman (84) investigated the relationships between semantic sensitivity ("an awareness that words have more than one meaning and the particular meaning implied varies with the context") and reading achievement as indicated by scores on the Reading Subtest of the California Achievement Test and the Gates MacGinitie Survey. With his 102 sixth grade Ss, he found high correlations (.72) between the Semantic Sensitivity Test and both achievement tests.

A study of children's thoughts about language in its spoken and written forms was reported by Downing (23). Three procedures were used to stimulate Ss' (thirteen British children approximately 5 years old) responses to concepts involved in reading and writing: (1) An open-ended interview, (2) an interview plus concrete stimuli such as color photographs, books, and toy busses, and (3) the first two plus a "yes/no game" using tape-recorded auditory stimuli such as non-human noises, phonemes, words, etc. The three research approaches were found to produce complementary findings. Young children find it difficult to understand the purpose of written language features. Furthermore, they have a special difficulty with abstract terms. The categories "word" and "sound" are very poorly understood. Downing suggests that children need to be helped to understand the purpose and nature of the written form of language.

Two experiments conducted by Schaeffer and Wallace (76) attempted to demonstrate the effects of semantic similarity on word-meanings judgments. Ten paid college students were the Ss in each experiment. In the first experiment, Ss reviewed pairs of words using a tachistoscope. If both or neither of the words were living or nonliving, Ss pressed a "same" key; if each word belonged to a different category the "different" key was pressed. Response latency in milliseconds was measured. Subcategories within the "living" category were mammal and flower; in the "non-living" category the sub-categories were metal and fabric. Results showed that mammal-flower and metal-fabric "same" judgments were more difficult than any sub-category "same" pairs, such as flower-flower.

In the second experiment, four category labels (bird, mammal, fruit, and grain) and twenty item-words, five in each category (duck, fox, plum, rice, etc.), were used. Subjects viewed a pair of category words followed by an item word, and then pressed a button indicating the appropriate category. Results showed that semantically similar labels (bird-mammal) were harder to differentiate than were semantically dissimilar ones (mammal-grain).

In both experiments, a repeated measures analysis of variance showed the average difference scores at the .001 level. The authors concluded that semantic similarity facilitates the judgment of meaning equivalence, but hinders the judgment of meaning difference.

Di Vesta and Ingersoll (22) investigated the effects of cue-oriented instructions, redundancy, and attribute saliency on concept identification in a verbal task. The Ss were 130 pupils in grades 4 through 6. They concluded that differences in concept identification related to the kinds of information a S is provided about the attributes in such a task, to the quantity of relevant attributes upon which the formation of the concept depends, and to the saliency of the attributes of the concept.

4. *Comprehension.* A study was done by Rystrom (72) to determine the extent to which his comprehension model accurately reflects the comprehension process. His model, defined in an earlier report, consists of a 6-skill definition of reading comprehension: vocabulary, syntax, item recall, item sequence, interpretation, and evaluation. Rystrom constructed a comprehension test containing fifty-seven items including some of each of the six different types he identified in his model. The number of questions of each type varied greatly: three evaluation questions to seven vocabulary, and fourteen item-recall questions, for example. All were multiple-choice questions about Slobodkina's *Caps for Sale,* written at about a second-grade level.

Each of his 169 fourth-grade Ss was asked to read the story as many times as he wished and was then given the test (Rystrom Reading Comprehension Test—RRCT). Results indicated that the test was reliable (.84) correlating odd- and even-numbered items, and that it correlated highly with the California Reading Test (.86). He found that Ss had least difficulty with item recall questions and most difficulty with item sequence questions. He also found that item recall was the best single predictor of reading comprehension.

Rystrom concludes that the RRCT appears to be both highly reliable and valid, that the six factors (skills) included do define comprehension, and "the most important single comprehension skill" appears to be recall of specific facts. These conclusions seem very tenuous in light of several observations:

1. *Only one grade-level of children was tested.*
2. *Only one story was used as a comprehension basis.*
3. *Types of questions asked consisted only of the six included in his model. Cause-effect, appreciation, etc., were omitted.*
4. *Very few items of the inferential or evaluative types were involved.*
5. *All questions were multiple-choice items which lend themselves more readily to recall than to more complex levels of comprehension.*

Finally, one wonders what is meant by "the most important single comprehension skill." If recall of details is the type of comprehension children are most frequently taught, it should not be surprising that it is in this area that they perform best—particularly on multiple-choice tests.

Mickelson (57) addressed herself to the effect of meaningfulness on children's verbal learning. To do this, she gave sixty-four randomly selected 9-year-olds a list of high-meaningful and low-meaningful words. A table of meaningfulness had been developed in an earlier study by her. A free-learning, free-recall design was used wherein Ss were asked to learn as many words as they could in 3 minutes and then recall them. A t-test of the difference between means showed that the high-meaningful items were learned significantly more readily (.01) than the low-meaningful words. One of her suggestions has considerable merit in reference to reading instructions:

> Meaningfulness, in fact, as defined in this study might be thought of as a basis of determining individual differences in children's verbal facility and as such, the starting point for individualizing reading instruction.

At the high school level, Andresen (1) sought to test the development of literary appreciation through teaching the ability to evaluate the profundity of literature. Subjects (N = 266) were divided into experimental and control groups. Experimental Ss were taught to evaluate the profundity of themes in literature in light of a Profundity Scale of five levels: physical (least profound), mental, moral, psychological, and philosophical (most profound). Control Ss were taught to evaluate literature through conventional informal discussion. Using two tests of literature interpretation, Andresen found that experimental Ss showed greater gains than control Ss, though only experimental girls achieved significant gains (.07). These findings indicate either a revision of the Profundity Scale or further research in its use is warranted before its dissemination as a tool for the development of comprehension and appreciation of literature. Andresen recognizes it as a pilot attempt to examine one phase of literacy appreciation.

Investigating reading comprehension in a different way, Zeman (99) sought to describe the relationship between measured reading comprehension and basic sentence types and sentence structural patterns used in the writings of second and third grade children. Reading comprehension scores were obtained from the Science Research Associates (SRA) achievement series; basic sentence types were the simple, complex, compound, and compound-complex and the ten basic sentence structural patterns were those identified by Roberts. To ascertain writing skills, the 192 second grade and 220 third grade Ss were asked to write an ending to an unfinished story. Zeman found that above-average readers used more compound sentences (.05). No causal relationships were suggested.

To help eleventh graders establish a framework for reading a story, Smith and Hesse (81) prepared an oral, tape-recorded "cognitive organizer." The cognitive organizer was constructed to give the S an organizational framework for the selection to be read, without disclosing significant details, inferences, or the main idea. Experimental Ss consisted of 169 eleventh grade pupils (98 designated good readers on the basis of the Test of Academic Progress (TAP) and 71 designated poor readers). There were 171 Ss in the control group (101 good readers and 70 poor readers). Both groups read the same selection, responded to an interest inventory, and completed a 7-item comprehension test. Prior to reading the selection, the experimental group listened to the 3-minute cognitive organizer.

Four variables were analyzed: detail recall, inference, main idea, and attitude. The investigators found, using multivariate analysis of variance, that good readers scored significantly higher (.0001) than poor readers on all variables. However, with exception of attitude, there were no significant differences between the experimental and control groups. The authors conclude that the use of a cognitive organizer seemed to have a positive effect on the attitudes of poor readers but was of dubious value, except possibly for developing positive attitudes, for good readers. Further, poor readers seemed to be helped by the cognitive organizer in determining main ideas, which was measured by one item on the comprehension test.

Since attitude toward reading is one of the major hurdles to overcome with poor readers, it seems the use of cognitive organizers would be worthwhile with such students.

Culbertson (18) reported on two studies with widely varying audiences and message types, but which yielded parallel results on factors influencing the judgment and interpretation of quantities presented within articles about controversial topics. He found that reducing contextual magnitude increased the apparent size of core quantities, where contextual relevance was high.

5. *Phonology*. The use of synthetic words or nonsense syllables is becoming increasingly widespread in reading research concerned with children's acquisition of letter-sound relationships. Hayden and Loud (39) devised a study to determine empirically the pronunciability values of two hundred nonsense syllables. The nonsense syllables were constructed following the guidelines of Hilgard. Thirteen percent of the syllables had been rated in an earlier study and the correlation between present and past ratings was established.

Undergraduate Ss (N = 201) rated each of the syllables on a 9-point scale, with a low value representing ease of pronunciability. The product-moment correlation between this set of ratings and the earlier one was .99. The resulting table of two hundred nonsense syllables with their mature-reader pronunciability ratings should prove a useful source for researchers interested in acquisition of grapheme-phoneme correspondences.

6. *Readability*. Whalen (92) conducted a study to test the assumption that the *Bond Street Readers*, designed for use with disadvantaged children, would be "easier" in terms of readability than the Harper and Row Basic Reading Program materials. Although Whalen states that one might expect the pacing from pre-primer to primer to first reader to be slower and more repetitive, and that fewer and more "regular" words would be introduced, he hypothesized that the *Bond Street Reader* would, in fact, be more "difficult."

Two pre-primers, a primer, and first reader from each series were compared. Four variables were selected to gauge the readability of these materials: (1) number of irregular words (words which do not follow predictable or highly frequent letter-sound correspondence patterns, such as *any, debt,* and *hour),* (2) number of different words, (3) average sentence length, and (4) average word length. Chi-square tests were performed for irregular words and different words, while two-tailed t-tests were made of average word and sentence length.

The results of the statistical tests show significant differences between the two sets of materials, and that the *Bond Street Readers* are the more "difficult" particularly at the pre-primer level. Whalen naively concludes that, "In view of the fact that culturally disadvantaged children are believed to have language disabilities, the choice of the Bond Street series for their use is questionable." Perhaps the investigator is unaware of the considerable evidence that disadvantaged children possess language *differences* not *disabilities*. He has not examined oral language and dialectical differences of beginning readers using the two sets of materials. Such an investigation might produce much more valid information about the worth of the materials for standard or nonstandard speakers of English.

The plan of a study by Coke and Rothkopf (14) was to evaluate three algorithms for estimating the number of syllables in one hundred words of text. These were used to develop a computer program for determining a Flesch Reading Ease Score. The algorithms were based on the number of vowels (letters) per word, consonants (letters) per word, and letters per word. Using Pearson product-moment correlations with man-made syllable counts, they found the higher correlation with vowel letters (.92). The authors suggest that the use of computers in determining readability would allow more sample passages to be included, thus reducing the probability of miscalculation.

C. Perception

1. *Visual Perception and Processing.* Bonsall and Dornbush (6) investigated the relationships between reading ability and the "developmental stage of the child as a function of task difficulty, exposure, duration, and meaningfulness of visually presented material." Sixty Ss from grades 2, 4, and 6 participated. Subjects were presented pairs of words containing vowel clusters for different tachistoscopic exposure durations. They responded "same" or "different" if the vowel pairs were in the same or a reversed sequence. Some words were real (thus meaningful); others plausible or implausible synthetic words. Results indicated that younger Ss' performances might be limited by stimulus conditions, especially rapid duration, more than by reading ability as measured by standardized tests. The authors suggest that in beginning reading, the emphasis should be on accurate processing rather than reading speed.

The purpose of Brooks' (9) experiment was to extend the generality of conflict between reading and visualization he had discussed in a previous article. Subjects were given sentences describing a set of spatial relations: "The camel is to the left of the doghouse and below the sandbox." Ten college students comprised the sample. The results showed that Ss' visualization of the spatial relations was interrupted by reading the sentence, but not by listening to it.

Dick (21) examined the relation between information available in visual stimuli and information required in report. He reported that altering the form of redundancy affected accuracy of report, but did not systematically influence the amount of information lost. Spatial and color information were lost, but class was not.

In a study of potential importance for pre-reading and beginning reading instruction, Jensen and King (44) compared the effectiveness of three methods of visual-motor discrimination training (tracing, rearranging, and matching) on learning to read words. Word similarity and sex differences were also studied. The 126 kindergarten Ss were grouped according to method of training, and by sex and similarity or dissimilarity of words within group. The authors found that on the reading tests no significant differences resulted from the three different training methods. "No one method of visual-motor discrimination appears to be best as preparation for teaching words to all pupils." No significant sex difference was found, and there were no significant differences between similar and dissimilar words. Jensen and King caution readers to be skeptical of theories and programs which claim superior reading performance due to certain types of visual-motor discrimination training.

Several studies were concerned with the use of tachistoscopic devices. Winnick and Danials (97) conducted two experiments concerned with the effect of set on tachistoscopic thresholds. They found a sensitization effect accruing from previous viewing of printed words, but not where verbal responses had been evoked by pictures or definitions. As a study concerned with direction of eye movements during the occurrence of perceptual error, Mandes (54) reports that errors in the perception of a distinguishing feature are associated with initial eye movements to the opposite field of the distinguishing feature. In a study of the effect of retinal orientation on the recognition of novel and familiar shapes, Steinfeld (83) found that in conditions of 45°, 90°, and 180° retinal disorientation, such recognition did not significantly differ from the zero-level degree. At extensive degrees of disorientation, recognition of lengthy words did deteriorate, but these deficits were reduced by increasing exposure time. In Murch's study (54) words were presented tachistoscopically with different inter-stimulus intervals (ISI). As the ISI decreased from 1,000 to 50 millisecond, a decrement in performance was observed and explained as the result of a reduction in scanning time.

2. *Recognition and Discrimination.* In the study conducted by Fudin (29), 6-letter nonsense arrays were tachistoscopically presented successively in the right and left visual fields to twenty Ss. Four different displacements from a central fixation point were used. Significantly greater recognition scores were obtained for material exposed in the right visual field at the first three displacements and for the average of all displacements.

In all cases, this higher recognition score was limited to the three letters in the left-array half. The authors found their results in conflict with the evidence used for Hebb's notion of a cell assembly.

Seymour (79) conducted three studies of latency response initiation for complex clarification of four word-shape configurations. The results showed, "delays in response initiation, attributable to discrepancy between the display components, or the order of repeating identities and congruence, or alterations in the processing demands imposed on the subject. . . ."

Sixty-two kindergarten pupils constituted the sample in Karraker and Doke's study (46) of discrimination of letters of the alphabet. Early-late and progressive-constant dimensions of the method of stimulus introduction were employed in the discrimination of the letters *b* and *d*. The four resulting conditions were: Early-Progressive (EP), Late-Progressive (LP), Early-Constant (EC), and Late-Constant (LC). The results of the study favored the EP method of presentation, with the LC method doing the poorest. The authors maintain that further research in needed before instructional programs based on these procedures can be developed.

In a slightly different vein, Thornburg and Fisher (87) studied the discrimination of 2-dimensional letters by children after they played with 2- or 3-dimensional letter forms. Thirty preschool children were selected for the study. Pre- and posttesting of 2-dimensional letter discrimination was done using twelve randomly selected letters. One group of Ss had no further experience with letters between the two tests, group II played with 2-dimensional letter forms (two cards of each of the 26 letters) for a 10-minute period twice a week for 4 weeks, and the third group had the same number of play sessions, but played with 3-dimensional plywood letters. The results showed no significant difference in recognition of printed letters related to form of play experience, though differences occurred between the control group and the two experimental groups. Play with either 2- or 3-dimensional letters appeared to reduce errors in 2-dimensional letter recognition (.05). In a second experiment 2- or 3-dimensional tasks were alternated with the result that while 3-dimensional task experience reduced the errors on a subsequent 2-dimensional task, 2-dimensional task experience did not seem to reduce errors on a subsequent 3-dimensional task. Thornburg and Fisher's study seems to have implications for early-school, pre-reading activities.

In an interesting investigation of verbal organization and the facilitation of serial learning, Levin (50) gave 112 fourth and fifth grade Ss a 12-item serial list of high frequency nouns to learn. In some cases the words were presented in context and in others in isolation. It was found that presentation of the words in sentences was not *per se* a sufficient effector of facilitated performance on the task. However, in conditions where sentences were used to relate three or more successive items to one another, facilitation was observed. This finding has import for instruction in vocabulary development.

Calfee (11) reported on a series of studies of serial recognition memory in normal and retarded children. He found performance to be relatively constant over a wide range of age and IQ, and that, with the exception of response biases and forgetting rate, recognition memory processes of retarded and normal children appeared to be identical with those of adults.

3. *Bimodality.* Gerjuoy and Winters (31) studied free recall in bimodal presentation with eighty educable mentally retarded (EMR) adolescents. Eighty Ss were given a free-recall task in which the stimuli (line drawing pictures of 40 high frequency words) were presented bimodally (visual plus oral presentation) to determine whether recall differed from unimodal (oral) presentation in a previous experiment. Free recall of bimodally presented material was better than free recall following unimodal presentation.

Gaines and Raskin (30) investigated the influence of visual and tactual training and test modalities on form recognition by children with learning disabilities. The procedures consisted of visual or tactual training with unfamiliar stimulus forms followed by recognition tests on these forms in either the same or the other modality. Forty-eight 8- to 10-year old Ss participated. Training modality was a significant main effect (.01), with vision found to be better than touch. Training-test interaction was also significant with the visual-visual sequence leading to the fewest errors. Gaines and Raskin's findings raise questions about the widespread practice in special education classes of using a quantity of tactual-kinesthetic teaching practices.

D. Effects of Selected Factors on Reading Behavior

The "Effects of . . ." subhead is a new one for the review and summary. The fact is, however, that each year a substantial number of studies are designed explicitly to examine the effects of certain factors upon reading behavior. Those studies are brought together under this subhead.

1. *Reinforcement.* The claim is often made that reinforcement procedures that are successful with middle class children do not work with disadvantaged children. To get some empirical evidence, Pipulski (34) attempted to determine whether lower and middle class children would respond differently to different forms of reinforcement. Forty-eight kindergarten children, equal numbers of boys and girls and equal numbers of lower class Negroes and middle class whites, served in the study. The Ss were assigned to one of three reinforcement conditions: (a) knowledge of results, where the correct response was simply supplied; (b) social reinforcement, where a correct response was followed by praise; or (c) material reinforcement, where each correct response was followed by a candy award and each incorrect response was followed by a buzzer. The task was to learn a list of six previously unknown words to a criterion of two consecutive correct anticipations of the entire list or a maximum of twenty trials. Results from the middle and lower class groups were analyzed separately because the two groups differed significantly in Peabody Picture Vocabulary Test scores ($\overline{X} = 85$ for the lower class group; $\overline{X} = 102$ for the middle class group). The analyses showed that the middle class Ss responded best to the social reinforcement condition, but there was no such overall effect with the lower class group. Lower class girls did, however, respond best to social reinforcement, while there were no differences for the boys. The author stressed the failure of the material reinforcement to produce superior results, particularly among the lower class Ss, in discussing his results. Yet he himself pointed out the distraction caused by the candy in the learning situation. Perhaps the tight structure and rapid pace of an anticipation learning setup do not provide a satisfactory context for immediate candy awards with kindergarten children. How can a kid keep his mind on his work when there is candy to be eaten?

Spence (82) addressed herself to the problem of the distracting effects of material reinforcers. She tested her hypothesis that children who receive candy rewards are distracted from the task at hand. One hundred lower class and one hundred middle class second and third graders were presented a verbal discrimination task. The Ss were systematically assigned to groups positively reinforced by a light signal, bean, or candy; half of the Ss in the first two groups (light-reward and bean-reward) were told they would be given candy for each reinforcer received. With both classes, the material reinforcers (candy) produced poorer performance than purely symbolic reinforcers (light). In the lower class groups, candy- and bean-reward Ss were inferior in performance to the light and bean Ss; light-reward Ss were slightly inferior to the latter. Results with the middle class Ss were similar, except that the light-reward group was more similar to the candy- and bean-reward groups. In general, the results support the hypothesis; but they are contrary to results which suggest that with conceptual tasks lower class, but not middle class, children perform best under material reward conditions. Spence suggests that (a) leaping to conclusions that lower class children are unresponsive to symbolic reinforcers may be premature, if not inappropriate, and (b) the relative effects of two types of reinforcers are highly dependent on the specific procedures employed.

Cotler (16) examined the effects of positive and negative reinforcement and test anxiety on the reading performance of fifth and sixth grade boys. While test anxiety was significantly related to both speed and accuracy of the reading performance, the reinforcement procedures employed had no effect.

2. *Stimulus Characteristics.* Postman (65) examined the effect of word frequency, as shown by the Lorge-Thorndike count, on acquisition and retention under a free-recall condition amounting to 2-second exposure of each of the twenty-four words on each list followed by a 2-minute written free-recall period. Retention and relearning were tested after 7 days. The main finding was that the frequency has a substantial effect on subsequent long-term retention.

Yuille and Paivio (98) examined the effects of content abstractness on the recall of connected discourse. Abstractness was found to affect total recall scores and the type of word recalled. Level of organization, i.e., syntactic versus random presentation of the content words, also affected recall and the number and type of errors.

Rystrom (71) studied the effect of different representations of letters of the alphabet, properly known as allographs, e.g., A a *a* *A* a, on children's ability to identify the letters. Using a test he devised, the Rystrom Multiple-Form Letter Discrimination Test, he examined the ability of children in kindergarten, first, second, and third grade to deal with allographs. The main conclusion was that many children apparently are asked to perform complex reading tasks before they have developed adequate letter-discrimination skills. Exercises to teach children to identify all of the allographs of a particular grapheme are suggested. Davis and Otto (20) examined elementary school pupils' word associations to individual letters of the alphabet. The salient finding was that by sixth grade, children's most frequent associations to individual letters are similar to adult's associations. The results are discussed in terms of consensuality of associations given by Ss with different levels of reading ability. Nelson, Bercov, and Leslie (60) reported on the effect of assigning identical responses to words sharing letters in various ordinal positions. Working with college students, they found that different ordinal positions within words are differently effective as concepts: acquisition was fastest when identical responses were assigned to identical initial letters, next for middle-letter assignment, and last for terminal-letter assignment. This change from the first to last to middle letter prominence usually demonstrated was attributed to a left-right processing of the concept formation aspect of the task involved in the study. The conclusion was that "the beginning-end-middle sequence effect is limited to situations in which the stimuli are processed and encoded as units."

To examine the effect of disorientation on the discrimination of print, Clay (13) tested the hypothesis that if objects are viewed repeatedly in only one orientation there will be an increase in the effects of relative disorientation on discrimination. Performance of eighty-seven children, between the ages of 5.0 and 8.0 years, was monitored over a 3-year period. The children were asked to read fifteen words in each of three treatments: normal, reversed, and inverted. Good readers were disturbed by changes in orientation, particularly the reversed condition; but little concern was shown by the poor readers. Changes in performance over age were also noted: reversed orientation was most troublesome at first, but inversions became more difficult as reading improved. The fact remains, though, that with increased age and progress in reading the children gained in ability to read disoriented print. Thus, while the data tend to support the hypothesis, data from older Ss, at least those who read well, might not.

Carver (12) examined the effect of "chunked" typography on reading rate and comprehension. Working with college students, the author presented reading materials in five variations of "chunked" formats; that is, he presented sentences that were spatially separated into small groups of meaningfully related words. In three carefully planned experiments, there was no evidence of any salutary effect of chunking. While this was contrary to results that have shown facilitation of free recall, rote memorization, and speeded comprehension by chunking, the author quite justifiably points out that those activities are quite different from normal reading by mature readers. The conclusion was that chunking "will probably not improve the reading efficiency of mature readers, reading at their normal rate, no matter what method is used to separate the material."

Corsini, Jacobus, and Leonard (15) showed that preschool children were better able to remember pictures than words in a recognition-memory task. This was interpreted as support for Bruner's suggestion that young children, whose dominant mode of representation is ikonic (imagery), should be able more successfully to encode pictorial than verbal-symbolic material.

3. *Organization of Materials.* The effect of preliminary questions and visual displays on learning from an oral presentation was studied by Dwyer (25). Working with college students, she presented a 2,000 word instructional unit on the physiology of the heart on audio tape. The tape was preceded by thirty-nine questions designed to direct attention to visual illustrations that accompanied the oral presentation. Five types of illustration, displayed on slides, were used with separate groups: names of parts and verbal descriptions of processes covered in the presentation; simple line drawings; detailed, shaded drawings; photos of a heart model; and photos of an actual heart. Presentation time was equal for all Ss, and all responded to a 78-item test. The group which viewed the simple line did best on the criterion test. The author concluded that (a) all types of visual illustrations are not equally effective, and (b) the questions did not serve to focus Ss' attention when the more realistic illustrations may, of course, have distracted the Ss from the critical aspects of the presentation.

Frase, Patrick, and Schumer (28) observed that studies suggest that frequent post-questions, rather than prequestions, appear to facilitate learning through reading. Yet they observed, too, that the effect is more apparent when the Ss are participating in an experiment than under more naturalistic circumstances. Consequently, they studied the effect of question position and frequency on learning through reading under three different levels of incentive. College students (N = 270) were assigned to a control group or to one of three incentive conditions—they were told they would get nothing, \$.03, or \$.10 for each correct response to test items—and to subgroups for pre- or post-questions, frequent or infrequent questions, and relevant or incidental test items. The main finding was that the relative advantage of post-questioning tended to decrease as motivation (incentive) increased because the performance of the pre-question and control groups improved as incentive increased. A major conclusion was that the inclusion of motivational factors complicates the attempts to understand the behaviors involved in learning from textual presentations.

In another study, Frase (27) examined the influence of "conceptual clustering" on the level and organization. Forty-eight factual-descriptive sentences were ordered by concepts, attributes, and randomization and presented to equal size subgroups of forty-two female adults; half were told the conceptual structure of the passage before reading and half were not. Subjects who knew the general structure did better on the recall task, but this advantage was most clear as learning progressed. Acquisition was better on the well-organized passages presumably because the Ss devoted more attention to organizing the poorly written material. A major suggestion was that instructional techniques which facilitate selection of conceptually related sentences during reading should facilitate learning. Color coding of conceptually related sentences was mentioned as one possible means for overcoming defects in some materials.

The plan of a study by Rothkopf and Thurner (69) was to examine the effect of frequently occurring key words in a reading assignment on Ss' subsequent verbal output. Very simply, the attempt was to determine whether Ss would use frequently occurring key words in essays written after exposure to the words in reading. Despite the fact that the high school students who served as Ss appeared to learn little of the substance presented, their post-reading output did come to resemble the input from reading. The authors speculated that students' ability to use textbook terminology—or, to use the authors' words, "to simulate the lexical statistics of a particular word swarm"—may not be good evidence of a basic understanding of the subject matter. Nevertheless, teachers may be prone to reward the student who uses the words despite his lack of conceptual understanding. The study introduces techniques and raises questions that will be of interest to both researchers and to subject matter teachers as they consider further the attributes of the academic snow job.

Hall (35) studied the effect of three sets of learning instructions on first and third graders' ability to determine whether or not words had previously been presented. Pre-presentation instructions were designed to (a) encourage associations to words on the acquisition list, (b) restrict associations, or (c) permit the Ss to select their own strategy. As expected, new words were falsely recognized if they were strong associates of the words in the acquisition list. The instructions that encouraged associations resulted in more false recognitions, which suggests that an instructor-imposed strategy may be effective.

Considerable evidence has accumulated to show that interrupted tasks are recalled and resumed more often than tasks previously completed. Ryan and Macmillan (70), however, pointed out some task characteristics in the previous studies that impose restrictions on drawing inferences for the classroom: (a) the tasks were primarily non-instructional and perceptual-motor in nature, (b) the tasks have generally been presented to individuals rather than groups, (c) Ss traditionally have been asked to choose resumption of a task completed or a task interrupted, and (d) criterion scores have been in terms of recall of tasks worked on rather than resultant learning. To study the effect of interruption in a more classroom-like application, they designed a study in which (a) an instructional sequence in social studies was employed,

(b) Ss were dealt with in groups, (c) students made a *resumption* rather than a repetition choice after interruption, and (d) the criterion test measured learning on the instructional task. In addition, high- and low-anxious fourth graders participated as Ss to permit examination of a possible interaction between test anxiety and task interruption. Analyses of the data revealed no task interruption effect on the learning of social studies materials, no test anxiety effect, and no interaction. The effect of task interruption was simply that it caused interrupted Ss to take longer to complete the task than uninterrupted Ss.

Nodine and Hardt (61) studied the effect of letter-position cues on learning to read words. They reported that "pre-reading children depend to a great extent on features of individual letters in making discriminations among words." Smith and Landy (80) reported support for their hypotheses that (a) more work is required in the vocalization of non-word-like than word-like letter strings, and (b) the number of correct reproductions correlated significantly with mean effort ratings for both types of words.

A study designed to explore the influence of similarity and review of interpolated learning materials on delayed recall was reported by Gillman (32). Analyses of results revealed a significant main effect for degree of similarity and a significant interaction, but no significant effect for the review variable. Jones and Brehm (45) reported that the persuasiveness of a one-sided communication is reduced when the Ss know there are two sides.

4. *Situational Factors.* Mumpower and Riggs (58) examined the effect of parental pressure on children's sight recognition of words. A pool of children who showed either over-achievement or under-achievement on the Word Accuracy subtest of the Gilmore Oral Reading Test was identified. Thirty-three pairs of the over- and under-achievers were matched on the basis of age, sex, grade, IQ, and mental age. Subjects ranged in age from 6 to 16 and in IQ from 51 to 125. The pairs of Ss were compared on the basis of ten factors described in their case histories. Differences were found for a cluster of factors that seemed to suggest that the over-achievers exerted much effort in school, presumably in response to parental pressure. The over-achievers showed signs of emotional disturbance related to school, presumably caused by the parental pressure. The over-achievers excelled only in word accuracy, not in comprehension. The authors suggested that by exerting pressure parents expose their children to emotional disturbance and drive them to excel in rote learning, possibly at the expense of more complex learning. The results are interesting, but the procedures are questionable and the inferences are strained.

Hultgren and Crewe (43) examined the effect of time in the retention of college text material in an attempt to clarify the role of cramming as compared to a study-review procedure in preparing for examinations. They reported that when the total amount of study time is controlled and when the time between original study and review is controlled, the retention performance of college athletes who cram does not differ from that of athletes who employ a study-review strategy. The authors suggest that while study-review may be the more practical strategy, cramming seems to pay off.

White and Aaron (93) reported a correspondence between the affective dimensions of semantic ratings and cognitive components expressed in a cloze procedure with fifth grade Ss. Harrison (38) reported that "words presented at a stimulus level just adequate for producing chance recognition, and associated with the abrupt onset of shock pairing, appear later in recall than words presented at the same level but associated with abrupt cessation of shock pairing." The results were interpreted as evidence of subliminal perception. Samuels (73) reported that college students' word-recognition speed was facilitated when associative connections between words in text-matched readers' word associations. That is, word recognition was more rapid when Ss had prior knowledge of what the word might be.

E. Associative Learning

> *The focus of all of the studies reviewed here is upon associative learning; a paired-associate learning task is employed in most, but not all, of the studies.*

1. *Stimulus Characteristics.* In several studies the influence of certain stimulus characteristics on associative learning was investigated. The studies are arbitrarily placed here rather than in the "Effects of . . ." section of the review.

Samuels and Wittrock (75) investigated the effect of strength of association between two words on learning to read the second word in a 2-word pair. All of the thirty-six first graders who served as Ss were able to read the first word, but none of them were able to read the second word in eight pairs of words when the experiment began. Word pairs in the control treatment had no associative connections, and increasing amounts of associative strength between adjective-noun pairs was conditioned for the three additional treatments. A word recognition test on the second word in each pair was given after word-association training. The main finding was equally facilitated by all three degrees of word-association strength: all three conditions were significantly better than the no-associations condition, but the three did not differ from each other. The authors suggest that the finding that even minimal word association training increases reading attainment has implications for teaching: new words ought to be introduced in a context of known words with greater than zero associative strength. Once again, the wisdom of a practice devised long ago by perceptive teachers is validated in a laboratory study.

McCutcheon and McDowell (55) sought an answer to the question: To what extent does intralist similarity effect beginning readers' (a) rate of acquisition, (b) subsequent recognition skills, and (c) tendencies toward generalized incorrect responses to other words. Working with kindergarten children they found that more trials were required to master a list of high similarity words than a list of low similarity words. On the other hand, Ss who learned the low similarity list made more incorrect identifications and more false generalizations when they were confronted with new but similar words. The data support prior arguments that teaching beginning reading with words that exhibit gross perceptual differences may engender a tendency to learn words on the basis of a single letter or some other readily apparent cue. Thus, the price for rapid acquisition appears to be inferior performance in subsequent word attack. If this is so, the price is far too high.

Kellas and Butterfield (47) reported that response pre-training facilitated the paired-associate learning of their third grade Ss, and that the magnitude of the facilitation was directly related to the rated pronunciability of the trigrams. These findings are in line with results previously obtained with adults.

That support for the hypothesis that the probability of a particular associate being involved in the encoding of a word presented in a meaningful context, and thus occurring as an implicit associate response to that word, is related to the "semantic acceptability" of the associate in the particular context was reported by Hall and Crown (36). General confirmation of the hypothesis that rhymes are normally easily learned because knowledge of the rhyming relation considerably restricts the range of response alternatives that need consideration for a given stimulus was reported by Bower and Bolton (8). Davidson and Dollinger (19) reported facilitation of second graders' paired-associate learning for sentences where an integral subject-object relationship could be specified between the paired items in some underlying deep structure form. Walker, Montague, and Wearing (88) reported that associability of the two items in the pair is a strong predictor of paired-associate acquisition differences with college students; but their data did not show definitively whether performance depends on the specific characteristics of the pairs.

2. *Stimulus Presentation.* The focus was on the manner of stimulus presentation in the studies that follow.

Samuels (74) assigned sixty kindergarten pupils to one of two discrimination training groups: (1) simultaneous training, in which a stimulus letter and response alternatives to be matched were presented together, and (2) successive training, in which the stimulus letter and response alternatives were presented by themselves. The purpose was to compare the effect of simultaneous versus successive discrimination training on transfer involving paired-associate learning, where the task was to learn the word response to a letter from the training task. Four similar letters—b, d, p, q—were the stimuli in both the discrimination and transfer tasks. The results support previous findings: when stimulus similarity is high, simultaneous discriminations are easier than successive discriminations. On the transfer task, however, the Ss who received the successive discrimination training did best. The latter finding is in conflict with an earlier finding of no difference between discrimination groups on a transfer task. Samuels suggested that successive training may be superior only when stimulus similarity is high.

Keppel and Zavortink (48) reported that college students who learned a paired-associate list under conditions of interference were more susceptible to unlearning from additional interpolated interference than students who learned under noninterference conditions. Hall and Kozloff (37) reported that the frequency of college students' false recognitions of the associates of early appearing words first increases and then declines as a function of the number of times their corresponding critical stimulus words had occurred. These results were explained in terms of perceived situational frequency. Bilsky and Evans (4) asked thirty-two mentally retarded Ss, age 12 to 19, to recall twenty words from four conceptual categories on four consecutive trials. The words were presented to one subgroup in random order for all four trials. To a second subgroup, the words were presented in categories on the first two trials and in random order on the last two trials. Both groups were divided into above- and below-median reading comprehension subgroups. Two main findings were that (a) organized presentation increased clustering on subsequent non-organized trials, and (b) above-median readers clustered more than below-median readers. The authors suggest that a tendency to cluster responses during free recall may reflect organizational processes that are central to the attainment of certain reading skills, like comprehension. The results support that suggestion as well as additional speculation that clustering may be facilitated by manipulating input organization. While there is evidence that increased organizational skills will transfer from one situation to another, the conditions necessary for long-term transfer must yet be identified.

3. *Mode of Presentation.* Here the focus was on the *mode*, rather than the *manner*, of stimulus presentation. Using a procedure similar to that in the Bilsky and Evans (4) study, Evans (26) examined the effects of mode of presentation on the associative clustering and recall performance of sixty mentally retarded adolescents. The Ss were divided into above- and below-mean groups on the basis of reading test scores, and a list of twenty words from four conceptual categories was presented via the auditory, visual, or auditory-visual modes to subgroups of equal size. The bimodal auditory-visual presentation had a facilitating effect on recall but not on clustering, and in this study reading level was not related to clustering or to recall. The nature of the relationship between clustering and reading performance, if any, apparently needs clarification.

Blue (5) examined the influence of mode of presentation, age, and intelligence on paired-associate learning. Subjects were thirty-six retarded high school age students, and seventy-two normal Ss, thirty-six of equal CA and thirty-six of equal MA. Equal numbers of Ss from each category learned a list of paired-associates with visual presentations of the stimulus portions and auditory, visual, or auditory-visual presentations of the response portions. Analyses of trials-to-criterion scores revealed no significant treatment effects or interactions. Performance of the normal Ss was significantly superior to that of both the CA and MA retardate peer groups. The author suggested that familiarity, meaningfulness, and/or pronunciability as well as modality need to be considered as relevant task variables in studies of the paired-associate learning of retarded Ss.

4. *Retardates' Performance.* In the studies that follow, the general purpose was to compare associative learning performance of normal children and retardates.

A group of normal children, CA 2 to 14 years, was MA-matched for contrast to groups of children with mongolism, encephalopathy, or retardation of unknown etiology by Sersen, Astrup, Floistad, and Wortis (78). The main purpose was to seek comparisons with certain observations made by Russian authors on the basis of CA rather than MA contrasts of retardates' and normals' performance. The concern was that the differences found may simply have been a reflection of lower intelligence rather than defects in a specific set of processes. Accordingly, a motor conditioning procedure commonly used in the Soviet Union and a word-association test were applied to the MA-matched groups. The motor conditioning task involved a response reversal sequence with a light signal-button pressing set up; and the word-association task involved the standard first association procedure. Differences were found between the normal and retardate group in motor conditioning performance; but, most relevant in the present context, no differences were found in the number of primitive word-association responses. The latter finding was explained in terms of the MA matching: if, as suggested, the word-association task merely amounted to an unstandardized intelligence test, then no differences should have been expected.

Proceeding from a prior observation that retardates have difficulty in " 'screening out' or extinguishing intruding responses" in paired-associate learning, Drew and Logan (24) compared the proportions of extra-list errors made by normal and retarded Ss as a function of associative strength of the list. Forty-eight normals and forty-eight retardates, age 10 to 12, were assigned to learn either a low or a high associative-strength list of fourteen pairs. The retardates imported a significantly greater proportion of responses under both levels of associative strength, but the proportion of imported responses did not differ as a function of associative strength. The authors suggest that instructional materials designed to minimize intrusions might improve retardates' performance on rote-learning tasks. Unfortunately, their data offer no clues as to how such materials might be structured.

Baumeister and Berry (2) attempted to assess the amount and locus of cue selection in the paired-associate learning of normal children and retardates. The retarded Ss ranged in age from 12 to 30, and the normals were in grades 3 to 6. The Ss learned a 5-item, Consonant-Vowel-Consonant (CVC)-stimulus with single digit response, list, and then responded to a transfer task in which they were asked to fill in the missing letters when only the initial, medial, or final letter was given. Analysis of the transfer data showed that the first letter is the most effective recall cue for both normals and retardates; both groups were equally prone to cue select in a left to right sequence. The authors suggest that retardates are likely to behave "normally" when their attention is directed to a relevant dimension. On the basis of another study, Baumeister, Berry, and Forehand (3) suggested that retardates are particularly limited in their capacity to observe the various stimulus dimensions presented in a verbal paired-associate task.

REFERENCES ON THE PSYCHOLOGY OF READING

1. **Andreson, Oliver**, "The Significance of Profundity in Literary Appreciation," *Reading Research Quarterly*, 5:(no. 1)100–117, Fall 1969, II.B.4.
2. **Baumeister, Alfred A.; Berry, Franklin M.**, "Single-Letter Cue Selection in the Paired-Associate Learning of Normal Children and Retardates," *Journal of Experimental Child Psychology*, 9:400–410, 1970, II.E.4.
3. **Baumeister, Alfred A.; Berry, Franklin M.; Forehand, Rex**, "Effects of Secondary Cues on Rote Verbal Learning of Retardates and Normal Children," *Journal of Comparative and Physiological Psychology*, 69:(no. 2)273–280, 1969, II.E.4.
4. **Bilsky, Linda; Evans, Ross A.**, "Use of Associative Clustering Technique in the Study of Reading Disability: Effects of List Organization," *American Journal of Mental Deficiency*, 74:(no. 6)771–776, 1970, II.E.2.
5. **Blue, C. Milton**, "Influence of Mode of Presentation, Age, and Intelligence on Paired-associates Learning," *American Journal of Mental Deficiency*, 74:(no. 4)527–531, January 1970, II.E.3.
6. **Bonsall, Cheryl; Dornbush, Rhea L.**, "Visual Perception and Reading Ability," *Journal of Educational Psychology*, 60:(no. 4)294–299, 1969, II.C.1.
7. **Bougere, Marguerite Bondy**, "Selected Factors in Oral Language Related to First-grade Reading Achievement," *Reading Research Quarterly*, 5:(no. 1)31–57, Fall 1969, II.B.1.
8. **Bower, Gordon H.; Bolton, Laura S.**, "Why Are Rhymes Easy to Learn?" *Journal of Experimental Psychology*, 82:(no. 3)453–461, 1969, II.E.1.
9. **Brooks, Lee R.**, "An Extension of the Conflict Between Visualization and Reading," *Quarterly Journal of Experimental Psychology*, 22:91–96, 1970, II.C.1.
10. **Burke, Carolyn L.; Goodman, Kenneth S.**, "When a Child Reads: A Psycholinguistic Analysis," *Elementary English*, 47:(no. 1)121–129, January 1970, II.B.1.
11. **Calfee, Robert C.**, "Short-Term Recognition Memory in Children," *Child Development*, 41:(no. 1)145–161, March 1970, II.C.2.
12. **Carver, Ronald P.**, "Effect of a 'Chunked' Typography on Reading Rate and Comprehension," *Journal of Applied Psychology*, 54:(no. 3)288–296, June 1970, II.D.2.
13. **Clay, Marie M.**, "An Increasing Effect of Disorientation on the Discrimination of Print: A Developmental Study," *Journal of Experimental Child Psychology*, 9:297–306, June 1970, II.D.2.
14. **Coke, Esther; Rothkopf, Ernst Z.**, "Note on a Simple Algorithm for a Computer-Produced Reading Ease Score," *Journal of Applied Psychology*, 54:(no. 3) 208–210, 1970, II.B.6.
15. **Corsini, David A.; Jacobus, Kenneth A.; Leonard, S. David**, "Recognition Memory of Preschool Children for Pictures and Words," *Psychonomic Science*, 16:(no. 4)192–193, 1969, II.D.2.
16. **Cotler, Sheldon**, "The Effects of Positive and Negative Reinforcement and Test Anxiety on the Reading Performance of Male Elementary School Children," *Genetic Psychology Monographs*, 81:31–49, February 1969, II.D.1.
17. **Craik, F. I. M.; Masani, P. A.**, "Age and Intelligence Differences in Coding and Retrieval of Word Lists," *British Journal of Psychology*, 60:(no. 3)315–319, 1969, II.A.1.
18. **Culbertson, Hugh M.**, "The Interpretation of Messages in Light of Contextual Magnitude and Relevance," *The Journal of Communication*, 20:32–50, March 1970, II.B.4.
19. **Davidson, Robert E.; Dollinger, Laurel E.**, "Syntactic Facilitation of Paired-Associate Learning: Deep Structure Variations," *Journal of Educational Psychology*, 60:(no. 6)434–438, 1969, II.E.1.
20. **Davis, Mary Lou; Otto, Wayne**, "Children's Word Association to Individual Letters," *Journal of Reading Behavior*, 1:(no. 4)10–21, Fall 1969, II.D.2.
21. **Dick, A. O.**, "Visual Processing and the Use of Redundant Information in Tachistoscopic Recognition," *Canadian Journal of Psychology*, 24:(no. 3) 133–141, 1970, II.C.1.
22. **Di Vesta, Francis J.; Ingersoll, Gary M.**, "Effect of Semantic Redundancy on Children's Identification of Verbal Concepts," *Journal of Experimental Psychology*, 82:(no. 2)360–365, 1969, II.B.3.
23. **Downing, John**, "Children's Concepts of Language in Learning to Read," *Educational Research*, 12:(no. 2)106–112, 1970, II.B.3.
24. **Drew, Clifford J.; Logan, Don R.**, "Extra-List Errors in Paired-Associate Learning by Retarded and Normal Children," *Psychological Reports*, 26:467–471, 1970, II.E.4.
25. **Dwyer, Francis M.**, "Effect of Questions on Visual Learning," *Perceptual and Motor Skills*, 30:51–54, 1970, II.D.3.
26. **Evans, Ross A.**, "Use of Associative Clustering Technique in the Study of Reading Disability: Effects of Presentation Mode," *American Journal of Mental Deficiency*, 74:(no. 6)765–800, January 1970, II.E.3.
27. **Frase, Lawrence T.**, "Paragraph Organization of Written Materials: The Influence of Conceptual Clustering Upon the Level of Organization of Recall," *Journal of Educational Psychology*, 60:(no. 5)394–401, 1969, II.D.3.
28. **Frase, Lawrence T.; Patrick, Edward; Schumer, Harry**, "Effect of Question Position and Frequency upon Learning from Text under Different Levels of Incentive," *Journal of Educational Psychology*, 61:(no. 1)52–56, 1970, II.D.3.
29. **Fudin, Robert**, "Recognition of Alphabetical Arrays Presented in the Right and Left Visual Fields," *Perceptual and Motor Skills*, 29:15–22, 1969, II.C.2.
30. **Gaines, Beverly J.; Raskin, Larry M.**, "Comparison of Cross-Modal and Intra-Modal Form Recognition in Children with Learning Disabilities," *Journal of Learning Disabilities*, 3:(no. 5)6–9, 1970, II.C.3.
31. **Gerjuoy, Irma R.; Winters, John J., Jr.**, "Subjective Organization by EMR Adolescents in Free Recall: Bimodal Presentation," *American Journal of Mental Deficiency*, 74:509–516, January 1970, II.C.3.
32. **Gillman, Stephen I.**, "Retroactive Inhibition in Meaningful Verbal Learning as a Function of Similarity and Review of Interpolated Material," *The Journal of General Psychology*, 82:51–56, 1970, II.D.3.
33. **Golden, Nancy E.; Steiner, Sharon R.**, "Auditory and Visual Functions in Good and Poor Readers," *Journal of Learning Disabilities*, 2:(no. 9)476–481, September 1969, II.A.3.
34. **Grob, James A.**, "Reading Rate and Study-Time Demands on Secondary Students," *Journal of Reading*, 13:(no. 4)285–288, 316, January 1970, II.A.6.
35. **Hall, James W.**, "Effects of Three Variations in Learning Instruction on Children's Word-Recognition Performance," *Journal of Educational Psychology*, 60:(no. 6)461–464, 1969, II.D.3.
36. **Hall, James W.; Crown, Irene**, "Associative Encoding of Words in Sentences," *Journal of Verbal Learning and Verbal Behavior*, 9:303–307, 1970, II.E.1.
37. **Hall, James W.; Kozloff, Edward E.**, "False Recognitions as a Function of Number of Presentations,"

The American Journal of Psychology, 83:(no. 2) 272–279, June 1970, II.E.2.

38. **Harrison, Robert H.**, "Effect of Subliminal Shock Conditioning on Recall," *Psychological Review*, 75: (no. 1) 19–29, 1970, II.D.4.

39. **Hayden, Benjamin S.; Loud, Lorraine**, "Some Norms for the Pronunciability of Nonsense Syllables," *Psychological Reports*, 25:415–418, 1969, II.B.5.

40. **Henderson, Norman B.; Butler, Bruce V.; Goffeney, Barbara**, "Effectiveness of the WISC and Bender-Gestalt Test in Predicting Arithmetic and Reading Achievement for White and Non-White Children," *Journal of Clinical Psychology*, 25:(no. 3)268–271, 1969, II.A.1.

41. **Hirst, Wilma E.**, "Entrance Age—A Predictor Variable for Academic Success?" *The Reading Teacher*, 23: (no. 6)547–555, March 1970, II.A.5.

42. **Huelsman, Charles B., Jr.**, "The WISC Subtest Syndrome for Disabled Readers," *Perceptual and Motor Skills*, 30:535–550, 1970, II.A.1.

43. **Hultgren, Dayton; Crewe, James**, "Time as a Factor in the Retention of College Text Material," *Journal of Reading Behavior*, 1:49–51, Fall 1969, II.D.4.

44. **Jensen, Norma J.; King, Ethel M.**, "Effects of Different Kinds of Visual-Motor Discrimination Training on Learning to Read Words," *Journal of Educational Psychology*, 61:(no. 2)90–96, 1970, II.C.2.

45. **Jones, Russell A.; Brehm, Jack W.**, "Persuasiveness of One- and Two-Sided Communications as a Function of Awareness There Are Two Sides," *Journal of Experimental Social Psychology*, 6:(no. 1)47–56, 1970, II.D.3.

46. **Karraker, R. J.; Doke, Larry A.**, "Errorless Discrimination of Alphabet Letters: Effects of Time and Method of Introducing Competing Stimuli," *The Journal of Experimental Education*, 38:(no. 4)29–35, 1970, II.C.2.

47. **Kellas, George; Butterfield, Earl C.**, "The Interaction of Pronunciability and Response Pretraining on the Paired-Associate Performance of Third-Grade Children," *Journal of Experimental Child Psychology*, 9:265–271, 1970, II.E.1.

48. **Keppel, Geoffrey; Zavortink, Bonnie**, "Retroactive Inhibition for Lists Learned under Interference Conditions," *Canadian Journal of Psychology*, 23: (no. 4)245–253, 1969, II.E.2.

49. **King, Paul T.; Dellande, William D.; Walter, Terry L.**, "The Prediction of Change in Grade Point Average from Initial Reading Rates," *Journal of Reading*, 13:(no. 3)215–218, 245–246, December 1969, II.A.6.

50. **Levin, Joel R.**, "Verbal Organizations and the Facilitation of Serial Learning" *Journal of Educational Psychology*, 61:(no. 2)110–117, 1970, II.C.2.

51. **Lingren, Ronald H.**, "Performance of Disabled and Normal Readers on the Bender-Gestalt, Auditory Discrimination Test, and Visual-Motor Matching," *Perceptual and Motor Skills*, 29:152–154, 1969, II.A.3.

52. **Lyle, J. G.**, "Certain Antenatal, Perinatal, and Developmental Variables and Reading Retardation in Middle-Class Boys," *Child Development*, 41:(no 2) 481–491, June 1970, II.A.5.

53. **Magoon, Jon; Cox, Richard C.**, "An Evaluation of the Screening Test of Academic Readiness," *Educational and Psychological Measurement*, 29:941–950, 1969, II.A.5.

54. **Mandes, Evans**, "Eye Movements and Perceptual Error: A Developmental Study," *Psychonomic Science*, 19:(no. 4)237–238, 1970, II.C.1.

55. **McCutcheon, Beth A.; McDowell, Eugene E.**, "Intralist Similarity and Acquisition and Generalization of Word Recognition," *The Reading Teacher*, 23:(no. 2)103–107, 115, November 1969, II.E.1.

56. **Michal-Smith, Harold; Morgenstern, Murry; Karp, Etta**, "Dyslexia in Four Siblings," *Journal of Learning Disabilities*, 3:(no. 4)185–192, April 1970, II.A.4.

57. **Michelson, Norma I.**, "Meaningfulness: A Critical Variable in Children's Verbal Learning," *The Reading Teacher*, 23:(no. 1)11–14, 1969, II.B.4.

58. **Mumpower, D. L.; Riggs, Sharon**, "Overachievement in Word Accuracy as a Result of Parental Pressure," *The Reading Teacher*, 23:(no. 8)741–747, May 1970, II.D.4.

59. **Murch, Gerald M.**, "Growth of a Percept as a Function of Interstimulus Interval," *Journal of Experimental Psychology*, 82:(no. 1)121–128, 1969, II.C.1.

60. **Nelson, Douglas L.; Bercov, Steven; Leslie Lauren**, "Ordinal Positions of Letters Within Words as Concepts: Effect of Assigning Identical Responses to Words Sharing Letters in Various Ordinal Positions," *Psychonomic Science* 18:(no. 4)201–202, 1970, II.D.2.

61. **Nodine, Calvin F.; Hardt, James V.**, "Role of Letter-Position Cues in Learning to Read Words," *Journal of Educational Psychology*, 61:(no. 1)10–15, 1970, II.D.3.

62. **O'Kelly, E.**, "A Method for Detecting Slow Learning Juniors," *Educational Research*, 12:(no. 2)135–139, February 1970, II.A.5.

63. **Parker, D. H. H.**, "Musical Perception and Backwardness in Reading," *Educational Research*, 12:(no. 3) 244–246, June 1970, II.A.3.

64. **Pipulski, John**, "Effects of Reinforcement on Word Recognition," *The Reading Teacher*, 23:(no. 6) 516–522, March 1970, II.D.1.

65. **Postman, Leo**, "Effects of Word Frequency on Acquisition and Retention Under Conditions of Free-Recall Learning," *Quarterly Journal of Experimental Psychology*, 22:185–195, 1970, II.D.2.

66. **Reed, James C.**, "Children's Figure Drawing—A Clue to Reading Progress," *The Reading Teacher*, 23: (no. 2)132–136, 143, November 1969, II.A.3.

67. **Rice, Donald B.**, "Learning Disabilities: An Investigation in Two Parts," *Journal of Learning Disabilities*, 3:(no. 3)149–155, March 1970, II.A.4.

68. **Roberts, Dennis M.; King, F. J.; Kropp, Russell P.**, "An Empirical Investigation of Ferguson's Theory of Human Abilities," *Canadian Journal of Psychology*, 23:(no. 4)254–267, 1969, II.A.6.

69. **Rothkopf, Ernst Z.; Thurner, Ronald D.**, "Effects of Written Instructional Material on the Statistical Structure of Test Essays," *Journal of Educational Psychology*, 61:(no. 2)83–89, 1970, II.D.3.

70. **Ryan, Frank L.; Macmillan, Donald L.**, "Effects of Interruption During Social Studies Instruction on Learning Effectiveness and Efficiency," *Journal of Educational Psychology*, 61:(no. 1)27–32, 1970, II.D.3.

71. **Rystrom, Richard**, "Evaluating Letter Discrimination Problems in the Primary Grades," *Journal of Reading Behavior*, 1:38–48, Fall 1969, II.D.2.

72. **Rystrom, Richard R.**, "Toward Defining Comprehension: A Second Report," *Journal of Reading Behavior*, 2:(no. 2)144–157, 1970, II.B.4.

73. **Samuels, S. Jay**, "Effect of Word Associations on the Recognition of Flashed Words," *Journal of Educational Psychology*, 60:(no. 3)97–102, 1969, II.D.4.

74. **Samuels, S. Jay**, "Effect of Simultaneous Versus Successive Discrimination Training of Paired-Associate Learning," *Journal of Educational Psychology*, 60:(no. 1)46–48, 1969, II.E.2.

75. **Samuels, S. Jay.; Wittrock, Merlin C.**, "Word-Association Strength and Learning to Read," *Journal of Educational Psychology*, 60:(no. 3)248–252, 1969, II.E.1.

76. **Schaeffer, Benson; Wallace, Richard**, "Semantic Similarity and the Comparison of Word Meanings," *Journal of Experimental Psychology*, 82:(no. 2) 343–346, 1969, II.B.3.

77. **Seraflca, Felicisima C.; Sigel, Irving E.,** "Styles of Categorization and Reading Disability," *Journal of Reading Behavior*, 2:(no. 2)105–113, Spring 1970, II.A.2.

78. **Sersen, Eugene A.; Astrup, Christian; Floistad, Ivar; Wortis, Joseph,** "Motor Conditional Reflexes and Word Associations in Retarded Children," *American Journal of Mental Deficiency*, 74:495–501, January 1970, II.E.4.

79. **Seymour, Philip H. K.,** "Representational Processes in Comprehension of Printed Words," *British Journal of Psychology*, 61:207–215, May 1970, II.C.2.

80. **Smith, Olin W.; Landy, Frank,** "Note on the Visual Recognition of Words," *Perceptual and Motor Skills*, 29:83–86, 1969, II.D.3.

81. **Smith, Richard J.; Hesse, Karl D.,** "The Effects of Pre-Reading Assistance on the Comprehension and Attitudes of Good and Poor Readers," *Research in the Teaching of English*, 3:(no. 2)166–177, Fall 1969, II.B.4.

82. **Spence, Janet Taylor,** "The Distracting Effects of Material Reinforcers in the Discrimination Learning of Lower- and Middle-Class Children," *Child Development*, 41:103–111, 1970, II.D.1.

83. **Steinfeld, George J.,** "The Effect of Retinal Orientation on the Recognition of Novel and Familiar Shapes," *The Journal of General Psychology*, 82:223–239, 1970, II.C.1.

84. **Svedman, Shirley,** "Semantic Sensitivity and Reading Achievement," *The Reading Teacher*, 23:(no. 7) 640–646, 1970, II.B.3.

85. **Tatham, Susan Masland,** "Reading Comprehension of Materials Written with Select Oral Language Patterns: A Study at Grades Two and Four," *Reading Research Quarterly*, 5:(no. 3)402–425, Spring 1970, II.B.1.

86. **Thompson, Mary R.; Whitehill, Richard P.,** "Relationships Between Reading Flexibility and Speed Gains," *The Journal of Educational Research*, 90:(no. 5) 213–215, January 1970, II.A.6.

87. **Thornburg, Kathy R.; Fisher, Virginia Lee,** "Discrimination of 2-D Letters by Children After Play With 2- or 3-Dimensional Letter Forms," *Perceptual and Motor Skills*, 30:979–986, 1970, II.C.2.

88. **Walker, Clinton B.; Montague, William E.; Wearing, Alexander J.,** "Natural Language Associability in Paired-Associate Learning," *Journal of Experimental Psychology*, 84:(no. 2)264–267, 1970, II.E.1.

89. **Ward, Byron J.,** "Two Measures of Reading Readiness and First Grade Reading Achievement," *The Reading Teacher*, 23:(no. 7)637–639, April 1970, II.A.5.

90. **Weisberg, Robert W.,** "Sentence Processing Assessed Through Intrasentence Word Associations," *Journal of Experimental Psychology*, 82:(no. 2)332–338, 1969, II.B.2.

91. **Wells, Marion E.,** "Preschool Play Activities and Reading Achievement," *Journal of Learning Disabilities*, 3:(no. 4)215–219, April 1970, II.A.5.

92. **Whalen, Thomas E.,** "A Comparison of Language Factors in Primary Grades," *The Reading Teacher*, 23:(no. 6)565–570, 1970, II.B.6.

93. **White, William F.; Aaron, Robert,** "Non-Verbal Cues as Determinants of Reading Comprehension," *Journal of Reading Behavior*, 1:53–65, Fall 1969, II.D.4.

94. **Whitehill, Richard P.; Jipson, Janice,** "Differential Reading Program Performance of Extraverts and Introverts," *The Journal of Experimental Education*, 38:(no. 3)93–96, Spring 1970, II.A.2.

95. **Williams, Maurice; Weinstein, Edith; Blackwood, Ralph O.,** "An Analysis of Oral Language Compared With Reading Achievement," *Elementary English*, 47:394–396, March 1970, II.B.1.

96. **Williams, Richard P.,** "Predicting Comprehension Gains of College Freshmen," *Journal of Reading Behavior*, 2:(no. 2)166–172, Spring 1970, II.A.6.

97. **Winnick, Wilma A.; Danials, Stephen A.,** "Two Kinds of Response Priming in Tachistoscopic Recognition," *Journal of Experimental Psychology*, 84:(no. 1)74–81, 1970, II.C.1.

98. **Yuille, John C.; Paivio, Allan,** "Abstractness and Recall of Connected Discourse," *Journal of Experimental Psychology*, 82:(no. 3)467–471, 1969, II.D.2.

99. **Zeman, Samuel S.,** "Reading Comprehension and Writing of Second and Third Graders," *The Reading Teacher*, 23:(no. 2)144–150, November 1969, II.B.4.

III. THE PHYSIOLOGY OF READING

A. General Health and Reading

To investigate the relationships between general physical health and reading disability, Morris (4) compared four hundred poor readers picked randomly from a St. Louis reading clinic's clientele with four hundred children of similar age, intelligence, and socioeconomic background in terms of childhood disease, surgical operations, and various specific physical defects. Most childhood diseases were nearly double for the experimental group, with mumps, measles, and chicken pox the most common. Surgical operations, too, were significantly more common among the poor readers. Of particular interest was the fact that, while auditory defects occurred more than five times as often among the poor readers, visual handicaps within the two groups were equal. It would seem that the implications here for instructional programs relying heavily upon phonetic approaches are obvious. The only area in which the control group exceeded the experimentals was in dental defects. Though the author had no ready explanation for this, he hypothesized that perhaps the experimental group's generally poorer health put them more often under medical care, thus encouraging more frequent dental examinations and subsequent treatment.

B. Sensory Deficit and Reading

1. *Deafness and Reading.* Hartung (2) compared thirty deaf children with thirty normally hearing children matched for chronological age and normal intelligence. To separate visual perceptual ability from code-knowledge, both Greek and English letters were used to produce pronounceable (phonologically regular) and unpronounceable (phonologically inconsistent or irregular) trigrams (a) in which the Ss noted the absence or presence of specific symbols, and (b) which the Ss were to totally reproduce. The results indicated that competence with the code, not visual perceptual skill, was what differentiated the two groups. At noting the presence or absence of specific symbols, the two groups did equally or almost equally as well, but at reproduction of the total trigrams, the hearing children did significantly better. For both groups, reading skill correlated highly with ability on the experimental tasks.

2. *Specific Learning Disability (SLD) and Visual and Auditory Processes.* The term specific learning disability is regularly used to describe the plight of children unable to normally perceive and process sensory experience even in the absence of gross neurological impairment. In order to appraise the psychosensory modality problems of such children, McGrady and Olson (3) administered a battery of thirteen intra- and inter-sensory tests to thirty-one 8- and 9-year-olds identified as SLD cases and a control group of sixty-eight normals matched for age, grade, sex, teacher influence, and socioeconomic factors. The larger control group was for the purposes of establishing normative data. Criteria for inclusion in the experimental group were (a) 90 IQ or better, (b) extreme difficulty in an academic area, and (c) being within the normal limits for vision, hearing, anxiety, and motor abilities. The tests consisted of identifying and/or matching such visual and auditory stimuli as words, pictures, or geometric figures

and words, familiar socially meaningful sounds, or pure tones. The SLD 8-year-olds had significantly more difficulty than their controls in every instance where visual symbols were utilized, and the 9-year-olds had greater difficulty regardless of the sense modality used. In particular, however, it was in terms of reaction time that the two groups differed, with the SLD's slower, particularly on verbal subtests. The authors concluded that the usual hypothesis of perceptual problems being at the root of specific learning disability remained unsupported, and that the experimental group's slow reaction time and weakness at verbal tasks were more meaningful avenues for further study.

C. Perceptual Style and Eye Movements

1. *Field Dependence/Independence and Reading.* Field-independent people, usually as identified via imbedded-figures tests, are supposedly more analytic in their perceptual styles than are field-dependents, and better able to note separate elements within a stimulus situation. To determine if these individuals would therefore attack word reading problems in a different way from field-dependents, Boersma and others (1) presented anagram tasks to field-independent and field-dependent Ss and examined their corneally reflected eye-movements. Field-independent Ss did indeed make more shifts in fixation per constant time-base, though successful performance *per se* did not differentiate the two groups. Also, males made more initial shifts per constant time-base than did females, and independent males attended to more letters per problem than did dependent males, though for females the reverse was true. In overall successful performance, there were no significant differences by sex.

Nadien and others (5), too, investigated field dependence and reading, but in interaction with weak or strong eye-dominance and S's mood. Thirty-five college-age students were identified in terms of (a) eye-dominance, (b) self-reported mood, and (c) field dependence, and then timed while reading a selection from the Nelson-Denny Reading Test. For Ss in a negative mood, field dependence was associated with weak eye dominance and poor comprehension, while for Ss in a positive mood, weak eye dominance was associated with slower reading and a strong correlation between reading speed and comprehension. The authors hypothesized that perhaps Ss who are habitually field dependent will, when they feel comfortable, act more independently and firmly than usual, and that this is why field dependence was related to weak ocular dominance only for Ss who were in a negative mood. No data differentiated by sex were reported.

2. *Familiarization Tasks and Subsequent Tachistoscopic Recognition, Paired-associates Learning, and Free Recall.* Winnick (6) compared the effects of two different familiarization procedures prior to (a) the securing of recognition thresholds for tachistoscopically presented words, (b) paired-associate learning, and (c) the free recall learning of presented words. Group A Ss were asked to make up as many new words as possible from the letters of six of the experimental words, a procedure hypothesized to be an aid to later tachistoscopic recognition, while group B was asked to search for associates to six of the experimental words, a procedure hypothesized to be an aid to later paired-associates learning and free recall. No significant differences by type of familiarization procedure were found, but the words used during the familiarization procedures were found (a) to have shorter tachistoscopic thresholds, (b) to take longer to learn by paired-associates, and (c) to differ only slightly in terms of criterion success for free recall learning.

REFERENCES ON THE PHYSIOLOGY OF READING

1. **Boersma, Frederic J.; Muir, Walter; Wilton, Keri; Barham, Richard,** "Eye Movements During Anagram Tasks," *Perceptual and Motor Skills,* 29:371–374, 1969, III.C.1.
2. **Hartung, Joseph E.,** "Visual Perceptual Skill, Reading Ability and the Young Deaf Child," *Exceptional Children,* 36:(no. 8)603–608, April 1970, III.B.1.
3. **McGrady, Harold J., Jr.; Olson, Don A.,** "Visual and Auditory Learning Processes in Normal Children and Children with Specific Learning Disabilities," *Exceptional Children,* 36:(no. 8)581–589, April 1970, III.B.2.
4. **Morris, John L.,** "The Link Between Reading Problems and Childhood Illness," *School and Community,* 56:(no. 6)23+, 26–27, February 1970, III.A.
5. **Nadien, Margot; Schaeffer, Donald S.; Schmeidler, G. R.,** "Mood as a Confounding Variable in Eye Dominance, Field Dependence, and Reading," *Perceptual and Motor Skills,* 29:277–278, 1969, III.C.1.
6. **Winnick, W. A.,** "Two Kinds of Familiarization in Tachistoscopic Recognition, Paired-Associate Learning, and Free Recall," *Psychonomic Science,* 16:(no. 4) 217–218, 1969, III.C.2.

IV. THE TEACHING OF READING

A. Assessing and Improving Reading Programs

Kingston and Cowart (26) replicated a study done by Gray in 1919 in which 250 content area teachers were asked to indicate the five purposes of reading which they thought most important for their students in preparing their course assignments. The purpose of the study was to determine any changes in teachers' perceptions between 1919 and 1969. The twenty-nine statements constructed by Gray were administered to 226 high school teachers. A rho coefficient was calculated and the relationship between the importance assigned to the various uses of reading by the two groups was found to be +.59. Further analyses indicated that modern teachers were more concerned with broader cognitive goals such as thought, reasoning, and the extension of knowledge. Of the various subject matters compared, modern language teachers showed the most consistency over the years. This study should be of interest to curriculum developers who view changes in teacher attitudes and behavior as major factors in educational improvement.

The questioning practices of primary grades teachers was the focus of a study by Bartolome (3). The questions asked by teachers were categorized and examined in relation to their objectives for primary reading instruction. The reading classes of eighteen teachers, six each in grades 1, 2, and 3, were observed by fourth-year education students. The questions used in the classes were categorized as requiring either memory, translation, interpretation, application, analysis, synthesis, or evaluation responses. The total number of questions recorded by the observers was 7,476. Of these, 47.54 percent were classified as memory questions. An important but not surprising finding is that the teachers tended to use lower-level questions for the slower students and higher-level questions for their more able students. Although the objectives frequently stated by the teachers involved analysis and application level thinking, the questions they frequently asked required only memory. The author concludes, as others have, that the need to develop higher-level intellectual skills among students is generally recognized as a worthy goal, but is seldom accomplished in the primary grades.

Two studies evaluated in-service programs designed to improve reading instruction. Smith, Bragstad, and Hesse (44) planned and conducted a half day in-service program to help high school teachers incorporate reading instruction into their content area teaching. A 5-item inventory constructed by the authors was administered to the participating teachers at the conclusion of the program. The responses to the inventory were scored and analyzed according to the various content areas represented. On the basis of the analyses the in-service program was judged to be effective in changing teacher attitudes and imparting knowledge. Three characteristics of the program were credited for much of the program's success: (1) reading was presented as behavior that can occur at various cognitive and affective levels, (2) faculty members who were already teaching reading in their classes were featured, and (3) specific instructional practices were suggested to the teachers.

Kasdon and Kelly (23) evaluated the effects of an in-service education program designed to help teachers select reading materials at appropriate instructional levels for their pupils. The focus of the program was practice administering informal reading inventories. Three hypotheses were tested: (1) an in-service education program designed to train teachers to determine pupils' instructional reading levels via simulation-type experiences increases teachers' awareness of pupils' instructional levels, (2) teachers who participate in the program before organizing classroom reading groups demonstrate a greater awareness of pupils' instructional reading levels than teachers who experience a similar program later in the year, and (3) primary grades teachers are more aware of the instructional reading levels of pupils than intermediate grades teachers. The Ss for the study were three stratified random samples of thirty-two teachers each from grades 2 to 5. One pupil was selected at random from each teacher's classroom and given the Standard Reading Inventory by McCracken. The difference between the pupil's score and the readability level of the basal reader the student was using in the classroom was used as the measure of teacher awareness. Analysis of variance was employed to test the three hypotheses. The results indicated that (1) a pure research design can be used to evaluate in-service programs, (2) simulation is an effective in-service technique, (3) in-service education is more effective when given before school begins, and (4) primary teachers do not demonstrate a greater awareness of the reading levels of their students than intermediate grades teachers.

Farr, Laffey, and Brown (16) conducted a questionnaire survey of all junior and senior high schools in the state of Indiana to gather information needed to develop training programs, improve certification requirements, and assess the school districts' needs for aid. The number of schools responding was 550 or 83 percent of those contacted. The responses to the questionnaire showed among other things that English teachers are usually assigned the job of teaching reading in secondary schools in spite of their lack of qualifications to accept this responsibility. Other valuable, but not unexpected findings are reported to show that the state of reading instruction in the secondary school is still far from satisfactory.

One study in this year's review addressed itself to the opinions and practices of kindergarten teachers regarding the teaching of reading. La Conte (28) surveyed 573 kindergarten teachers to find their opinions about teaching reading in the kindergarten and the instructional practices and materials they employed. An analysis of the teachers' responses to the survey indicated that (1) kindergarten teachers generally do not support the formal teaching of reading in kindergarten, (2) in most classrooms reading skills are taught occasionally, but not in a planned, sequential program, and (3) instructional reading materials are being used in many kindergarten classrooms. The author concludes that teachers seem to be unaware that their classroom practices do not reflect their attitudes. Perhaps the inconsistency is more a matter of agreeing upon a definition for reading.

B. Instructional Materials

Meisel and Glass (33) compared the content of the 1956 edition of the Scott Foresman "New Days and Deeds" basal reader with the content of the books fifth graders in one school voluntarily withdrew from the school library. In most cases, the reported high percentage of students indicating a preference for a particular interest area in voluntary reading choices did not match well with the percentage of times an interest area appeared in the basal text. In addition, many interest areas children indicated a preference for were not included in the basal reader. The implications of the study are limited because only one group of students and one basal reader were studied. The conclusion of the investigators that fault must be found with a basal text that neglects areas that children prefer needs a broader student subject base to be valid. Students in other geographic areas might register interests more in harmony with the content of that particular basal reader. Furthermore, the question of whether or not a basal reader should be expected to teach skills and appeal to students' interests is pertinent to the authors' conclusion.

Six basal reader series were examined by Stefflre (46) to determine to what extent the women characters in them were shown as workers and whether or not the working women in the books were distributed among the various occupations in the same way they are distributed in the actual labor market. The investigator found the discrepancies between the composition of the women's real labor force and the labor force as shown in the readers to be great and obvious. In discussing his findings the author did not fault the basal readers because of the discrepancy, but wisely observed that they do not have as one of their purposes the dissemination of accurate occupational information. The author concluded that since basal readers do not, counselors and teachers should give the child a more accurate picture of the place of work in the life of the modern woman.

Eight basal reading programs, grades 1 through 3, were analyzed by Schnayer and Robinson (41) to discover when each introduced the various phonic skills and the relative amounts of reinforcement each program suggested for the skills. The teacher's manual for each program was used exclusively for the tabulations that were made. Among the series examined, the authors found considerable inconsistencies regarding when each skill should be taught and how much reinforcement to provide for each skill. The authors list several limitations of the study, chief of which is their reliance upon the manual to answer the questions they asked.

Lam (29) attempted to determine the type of art preferred by ninety second-grade pupils when given a choice of the four major art styles used in illustrating primary reading textbooks: realistic, muted realistic, semi-abstract, and cartoon. A subproblem was the determination of the relationship of sex, intelligence, and socioeconomic status to pupil preferences. An instrument consisting of four sets of four pictures each was constructed and validated by judges to insure that it would test pupil preference for art style. A test retest reliability was conducted using tau. The range of tau was from .7564 to 1. The level of significance of these taus ranged from a Z of 5,043 to 6,667. All sixteen Z scores were significant beyond the .001 level of significance. Therefore, the instrument was considered to be highly reliable. The ninety Ss were tested individually and asked to indicate the picture they liked best, the one they liked second best, and the one they liked third best. The primary statistical treatments employed to evaluate the data were a chi-square test for independence, a chi-square Goodness of Fit test, and the Mann-Whitney U test. On the basis of the results obtained, the author concluded: (1) second-grade children are highly consistent in their rating of preferences for art style used in basal readers, (2) second-grade children rate realistic art, regardless of the picture content, as an overwhelming first choice, (3) second-grade children rate muted-realistic art, regardless of the picture content, as an overwhelming second choice, and (4) intelligence and socioeconomic status influence second-grade children's rating of cartoon and semi-abstract art styles only. Although the variables of intelligence and socioeconomic level may influence the third and fourth choices of pupils, they have no effect on first and second choices. Obviously, the major implication is the need for further research.

Reichard and Reid (39) tried to determine whether or not variable spacing and line verification had any effect upon the reading rate and comprehension of intermediate level educable mentally retarded (EMR) pupils. The Ss were eight classes of intermediate EMR pupils. A story at third-grade reading level was taken from the Reader's Digest "Skill Builders" series and duplicated in a manner to reflect four formats: (1) double-spaced verified right margins, (2) double-spaced unverified right margins, (3) single-spaced verified right margins, and (4) single-spaced unverified right margins. The assignment of format to each class was made on the basis of random selection. The eight teachers of the classes presented the story to their pupils, recorded the length of reading time and administered a comprehension test on the story. The data were analyzed by t-tests to compare mean reading time and comprehension scores for the different treatments. Subjects who read the double-spaced material received higher mean scores in both rate and comprehension than the Ss who read the single-spaced material. Higher mean comprehension scores were achieved by children reading double-spaced material with unverified right margins than by children reading double-spaced material with verified right margins.

Lockmiller and DiNello (30) investigated the relative effectiveness of "Words in Color" and Houghton Mifflin's first grade reader, "Up and Away," on the reading achievement of forty-eight second-grade retarded readers as measured by posttest scores on the McKee Inventory of Phonetic Skills and the California Reading Test. Analyses of the scores of the experimental and control group of twenty-four students each revealed no significant differences to favor one treatment over the other. The authors add their results to other studies which did not support the claims of superiority made for "Words in Color" over basal programs. This study was well designed, implemented, and reported. It is illustrative of the kind of research needed to support or refute claims of program superiority that are based more on wishful thinking than careful research.

Affleck (1) investigated the utility value of six phonic principles for intermediate grades reading materials. The vocabulary for testing the principles selected was the first 6,000 words in the *Thorndike-Lorge Teacher's Word Book*. A total of 19,400 syllables was classified, the vowel letter and sound identified, and the result of application of the appropriate principle recorded. These data were tabulated and two conclusions were drawn: (1) the percentage of utility value for vowel principles was found to be higher in this study than in previous studies and (2) the phonic principles investigated in this study had as high a utility value for intermediate grades level vocabulary as for primary grades vocabulary. The limitations and implications of the study are carefully listed by the author.

C. Developmental Aspects of Reading Instruction

1. *Pre-reading.* Several investigations reported during the preceding year dealt with the effects of visual-perception training and/or visual-motor training on various types of pupil performance. In one such study, Beck and Talkington (4) administered the Frostig-Horne visual perception exercises, 20 minutes a day for 7 months, to a sample of fifteen Headstart students (7 girls and 8 boys), while a comparable sample of fifteen Ss spent the 20-minute daily training periods in undescribed Headstart activities. At the end of the experimental period, the two groups were compared by applying Fisher t-tests to their mean difference scores from pre- and posttests on the Frostig Developmental Test of Visual Perception (FDTVP) and the Peabody Picture Vocabulary Test (PPVT). While the Ss who received the training in visual perception demonstrated greater mean gains in all FDTVP areas than did their counterparts, they were significantly superior in only Area IV— Position in Space ($p < .01$). Marked gains were observed when pre- and posttest performance on the PPVT were compared within groups; however, no significant differences were found when the mean difference scores on the PPVT for the two groups were tested. On the basis of the results, Beck and Talkington speculated that the IQ gains on the PPVT were attributable to the Headstart program in general rather than to the Frostig-Horne exercises in specific.

In a related study, Keim (24) hypothesized that kindergarten children who demonstrated visual-motor deficiencies on the Bender Visual-Motor Gestalt Test (Bender) would improve their performance on the PPVT, the Stanford-Binet Intelligence Scale (Binet), and the Metropolitan Readiness Tests (Metropolitan) as a result of visual-motor training provided by the Winter Haven Program (WHP). To test this hypothesis, Keim began with an initial sample of kindergarten children who met certain criteria. When the Koppitz Developmental Bender Scoring System was applied to their performances on the Bender, seventy-four Ss demonstrated visual-motor deficiencies. They were then assigned in equal numbers to an experimental group (E) and a control group (C–1). In addition, a second control group (C–2) was selected from the kindergarten children who had demonstrated no visual-motor difficulties. All three groups were equated on the basis of the PPVT, the Binet, and an undescribed Pre-Kindergarten Survey. During an undescribed training period, Groups C–1 and C–2 received an undescribed kindergarten program, while Group E received the same program plus the WHP. When the Ss' post-experimental period results on the Metropolitan, Binet, and PPVT were submitted to analysis of variance, no significant differences were found among the three groups with but one exception. The exception was that the three groups were significantly different on the Metropolitan Matching and Copying subtests. Additional testing of this finding revealed that only the Matching subtest scores provided a significant difference which favored Group C–2. Noting the limitations of the investigation, Keim concluded that the visual motor training provided by the WHP did not appear to accelerate the visual-motor development of Group E as measured by the instruments utilized in the study.

In a much more thorough and thoughtful effort than the two studies previously reviewed in this section Falik (15) set out to determine whether providing a group of kindergarten children with a specially designed curriculum, which emphasized (a) gross motor development, (b) eye-hand coordination, and (c) visualization patterns—form perception, spatial relationships, would have any effect on their readiness for reading at the end of kindergarten and on their reading skills in the middle of second grade. Falik selected and treated his Ss in the following manner: (a) the Anton Brenner Developmental Gestalt Test of School Readiness was administered to approximately ninety entering kindergarten children; (b) twenty Ss in the lower two-thirds on the Brenner Gestalt Test were randomly assigned to the experimental group and received the specially designed curriculum based on principles of perceptual motor development; (c) twenty-two comparable Ss were assigned to the control group and spent the year with an excellent kindergarten teacher, using the standard curriculum with a semi-structured experience designed to correspond in setting and general activity to the experience of the experimental group. Near the end of kindergarten the Ss took the Brenner Gestalt Test, The Metropolitan Readiness Tests, and the Schorr Test of Basic Perceptual Motor Development. A year and one-half later the Ss, sixteen experimental and seventeen control, took the reading section of the Metropolitan Achievement Test, Primary II Battery, Form B. When t-tests were applied to the scores on the end of kindergarten measures, no significant differences were found between the two groups. A similar finding existed when reading achievement at the middle of second grade was tested for the two groups. On the basis of his findings, Falik suggested that perceptual motor training might have utility with children who have developed problems in learning, but he was doubtful about the worthwhileness of such training for all children. More importantly, Falik raised some important questions about the measurement instruments used and about the ways in which they were administered and scored. Although his discussion of these issues raised doubts about some of his findings, he has provided some useful ideas for future researchers in this area.

Williams (50), in a well-detailed investigation, studied the relative influence of different training methods on the ability of kindergarten children to discriminate letter-like forms. To achieve this objective, the investigator conducted two experiments. Experiment 1 involved thirty-two children in the first month of kindergarten, while Experiment 2 employed thirty-two Ss in the last 2 months of kindergarten. In each experiment, three conditions were compared: (a) discrimination training where the comparison stimuli were quite different from the standard; (b) discrimination training where the comparison stimuli were transformations (rotations and reversals) of the standards; and (c) reproduction training where the S was asked to trace (twice) and copy (three or four times) each standard. It should be noted that the discrimination training consisted of delayed matching-to-sample tasks, in which the S was asked to identify the standerd after it was removed from view. Upon completion of the training in both experiments, a series of three tests was administered, and the same series was repeated 24 hours afterward. Among other things, Williams found in Experiment 1 that discrimination training in which the comparison stimuli were transformations was superior to discrimination training where the comparison stimuli were totally different forms and that reproduction training was not as effective as discrimination-with-transformations, but was as effective as the simple discrimination training. In Experiment 2, no significant differences among types of training were found. Williams concluded from the findings that more time should be devoted to discrimination training that involves comparison of letters with their transformations and that such training should be given early.

2. *Early Reading.* Durkin (14) described the achievement of thirty-six pre-first-grade Ss (20 boys and 16 girls) who participated in a 2-year language arts program which was based on the findings from her previous studies of children who learned to read at home. The Ss in this instance were volunteered for the program by their parents and were 4 years old by December 1 of the year the program began. One unique feature of the program was a 20-minute daily academic period during which the children worked on such things as reading, writing, letter identification, phonics, color names, and numerical identification. At the end of the first year, the Ss were assessed on their ability to identify numerals, letters, and words. They were evaluated on the same abilities at the end of the second year, as well as on their ability to identify letter sounds. In all cases the tests used were constructed by the investigator and had high content validity in terms of the program provided. Among other things, Durkin found that at the end of the second year the children could identify an average of 124 words, fifty lower-case and capital letters, forty-seven numerals, and fifteen letter sounds. Although the study was descriptive in nature, it provides the basis for comparative and longitudinal investigations Durkin will conduct as she follows the Ss through the elementary school.

To compare the relative effects of three kinds of discrimination training with relevant word forms on the performance of kindergarten children learning to read a word list, Jensen and King (22) randomly selected and assigned sixty girls to the following treatments: (a) tracing—similar words, (b) tracing—dissimilar words, (c) rearranging letters—similar words, (d) rearranging letters—dissimilar words, (e) matching—similiar words, (f) matching—dissimilar words. The same procedures were followed to assign sixty boys to the same treatment groups. In this instance, the similar words were feel, feet, fell, and full, while the dissimilar words were bark, city, jump, and shoe. After each of the four stimulus words for a particular group had been presented once in accordance with training procedures for that group, each S was tested on his reading of the words. Thus, each S was alternately trained and tested on the same words. The training tasks and reading tasks were repeated until twelve reading task trials had been completed or until the S had achieved a perfect score on four consecutive reading trials. Analysis of variance revealed a significant difference in mean number of correct responses among the groups. When the differences were compared by t-tests, tracing was significantly easier than the other two training methods, while matching was significantly easier than rearranging ($p < .001$). A 2-factor analysis of variance also revealed no significant differences for correct responses, errors, or omissions on the reading task. Finally, dissimilar words were significantly easier to learn to read than similar words ($p < .001$). The investigators concluded that their main findings did not support theories or practices which claim that children learn to read better by one or the other of the training methods studied.

The investigation by Jensen and King is a good example of the type of laboratory study that needs to be done in the area of early reading instruction. It is also a step in the direction of studying learning types. For readers interested in these important areas, the report of the investigation should be read in full.

3. *Initial Reading Instruction.* In an investigation designed to determine if Negro dialect is a cause of reading disability, Rystrom (40) tested three hypotheses: (a) Negro children can be taught to use features of white speech which do not occur in their native dialect; (b) knowledge of this additional dimension of dialect will have a positive and significant influence on word reading scores; (c) the use of phoneme-grapheme controlled readers will have a positive and significant influence on word reading scores. Four first-grade classes in an all-Negro rural school were included in the experiment: Class E-1 ($N = 24$) received dialect training and used a traditional basal reader; Class E-2 ($N = 23$) received dialect training and used a linguistic basal reader; Class C ($N = 23$) received no dialect training and used a traditional basal reader; Class E-3 ($N = 27$) received no dialect training and used a linguistic basal reader. Dialect training, which included work on the third person marker, terminal consonants, and terminal clusters, was allotted 20 minutes per day for 80 days. The non-dialect treatment was allotted the same amount of time and emphasized such language skill areas as vocabulary, sentence structure, listening, and comprehension. Efforts were made to control the teacher variable throughout the experiment. The Gates Word Reading Test and the Rystrom Dialect Test were given in January and again near the end of the school year. The reading sections of the Stanford Achievement Test were also administered as posttests to provide an index of general reading ability. When analysis of covariance was applied to the posttest results with the pretest results as the covariate, Rystrom rejected all three hypotheses. He concluded that there were still unanswered questions about the influence of dialect training on reading achievement. In fact, he pointed up the possibility of a negative influence, since he noted in his own data that children who received dialect training were able to read fewer words than the children who did not receive dialect training.

Warner (48) compared the mean gain scores on the Gates Primary Reading Test of six experimental first-grade classes ($N = 157$) that participated in a reading program where the pupils had a part in decision-making with the mean gain scores of six control classes ($N = 118$) that participated in a reading program where the teacher made the decisions. The Ss in the two groups were randomly selected and the groups were comparable in intelligence as measured by the Detroit Advanced First-Grade Intelligence Test. Teachers in the experimental and control classrooms were found to be similar in professional preparation and teaching experience. When t-tests were applied to the six mean gain scores derived from the Gates Test, the experimental group was found to be significantly superior in all cases ($p < .05$). Unfortunately the investigator did not give the pre- and posttest dates. Moreover, it was not clear from the report how the treatment conditions could be applied consistently across classrooms.

Among other things, Feldhusen, Lamb, and Feldhusen (17) found that a sample of first-grade children who received instruction from a basal series supplemented with programmed reading instruction were significantly superior ($p < .01$) in end of year reading achievement when compared with a sample of first-grade children who received instruction from a basal series only. In this instance, the criterion measures were the Word Knowledge, Word Discrimination, and Reading Comprehension subtest from the Metropolitan Achievement Test. It is to the investigators credit that they cited three limitations which prohibited the generalizability of the findings.

4. *Administrative Organization.* In a study to find the relative effects of graded and non-graded classrooms on the achievement of first- and second-grade boys and girls Brody (9) examined the scores of matched Ss on the Stanford Achievement Test. Mean scores for each of the five subtests of the instrument were analyzed by t-tests to determine any differences between the achievement of both first- and second-year students in graded and non-graded classrooms. Six of ten comparisons favored the students in non-graded classrooms at the .05 level of significance. The test scores were also analyzed to determine the effects of the different classroom organizations for high intelligence and low intelligence students. Students of low intelligence were found to benefit more from non-grading in the first grade and students with high intelligence in the second grade. Obviously, research is needed to determine whether or not one classroom organization is better than another to promote student achievement. However, this study did not control for the teacher and other classroom differences that could have been more responsible for the different achievement scores than the grading or non-grading organization. Therefore, the questions asked by the researchers remain unanswered.

5. *Reading Comprehension.* Using the Reading for Information and the Reading for Appreciation sections of the Bond, Balow, Hoyt, New Developmental Reading Tests—Intermediate Level, as the reading materials, Goudey (20) investigated the influence on reading comprehension when questions were asked before (directed) and after (non-directed) reading selections. Three hundred randomly selected fourth-graders were stratified into high, middle, and low reading achievement groups on the basis of Iowa Basic Skills results and were randomly assigned to treatments within groups. The investigator reported two major findings. First, the non-directed group scored significantly better than the directed group on the Reading for Information test. This finding was consistent at all three reading achievement levels. Second, there were no significant differences between the two groups on the Reading for Appreciation test. Goudey suggests, in closing, that directed reading, which is commonly stressed in professional textbooks, may not be as useful as assumed.

6. *Orthography and Reading.* Kresh (27) compared the reading achievement of 397 first-grade pupils, who were taught to read with the i.t.a. Scott Foresman series, with the reading achievement of 165 first-grade pupils, who were taught to read with the traditional Scott Foresman series. In this instance, reading achievement was measured with three subtests from the Metropolitan Achievement Test (taken in the traditional alphabet) and four subtests, including spelling, of the Stanford Achievement Test (taken in the alphabet in which the Ss were trained). When analysis of covariance was applied to the test data, the major results were: (a) when Metropolitan Readiness Test scores were used as a covariate, the i.t.a. Ss were significantly superior to the t. o. Ss on all the posttests; (b) when two phoneme scores from the Murphy Durrell Readiness Test were used as a covariate, the i.t.a. Ss were significantly superior on the four Stanford Achievement subtests and no differences were found on the Metropolitan; (c) when the results from the two readiness measures were used as the covariate, the i.t.a. Ss were significantly superior to the t.o. Ss on all the posttests. Although Kresh indicates that the Hawthorne effect was not controlled, it is commendable that the content of the instructional materials for the two groups was the same.

In a follow-up study, Downing and Latham (12) found that Ss from the original i.t.a. experiment were significantly superior to matched t.o. Ss on the NFER Sentence Reading Test 1, a group test of silent reading comprehension administered at the beginning of the fifth year of school, and on the NFER English Progress Test, administered at the end of the fifth year of school. The Sentence Reading study involved 156 Ss (78 i.t.a. and 78 t.o.), while the English Progress study involved 184 Ss (92 i.t.a. and 92 t.o.).

7. *Reading in the Content Fields.* The major purpose of a study by Smith and Burns (43) was to discover the relative effectiveness of three different pre-reading instructional practices on the attitudes and comprehension of high-ability ninth-graders. A ballad that appears in a widely used literature anthology was used for the experiment. Twenty, top-level English classes were randomly assigned to three treatments administered by the investigators: (1) general background information about this ballad and ballads in general before silent reading, (2) the same background information before an oral reading by a good reader, (3) the same background information before calling students' attention to words and word groups that needed careful attention while reading. A fourth group was used as a control and read the ballad with no pre-reading instruction. Attitudes toward the ballad were measured by a 6-item inventory constructed by the investigators, and comprehension was measured by asking students to select the main idea of the ballad from five choices. The scores on both instruments were analyzed by analyses of variance to determine any significant differences among groups. There were no significant differences among the four groups for attitudes, but the group which received the background information and was told to attend carefully to key words and word groups performed significantly better at the .01 level on the comprehension test. The authors concluded that the pre-reading instructional practices of teachers can make a difference in the interpretation of certain literary selections; however, the instructional practices typically employed may not be the most beneficial to a productive first reading.

Belcastro (5) looked at the effect of three levels of reading achievement on learning algebra through four techniques of programming. Eighth-graders ($N = 454$) were matched according to intelligence, reading scores, age, and arithmetic reasoning scores within sets of five. Each S in each set was then randomly assigned to one of the experimental groups or to the control group. The four experimental groups received programmed instruction in algebra for three, 43-minute periods. The four techniques were termed nonverbal inductive, nonverbal deductive, verbal inductive, and verbal deductive. The control group read non-mathematical material. Students of high, medium, and low reading ability were identified within each of the five groups, and analyses of variance were applied to the scores of the Ss on a 28-item posttest. Significant differences at the .05 level were found among the four experimental techniques for medium ability readers only. All ability levels within all of the four experimental groups significantly outperformed the control group on the posttest.

Keipe and Wood (25) compared the general English and reading achievement of fifty-nine sixth-graders who were randomly assigned to receive either a generative transform or a traditional approach to learning grammar. Both groups were pretested before instruction began and posttested after approximately 4 months of instruction with the reading and language arts subtests of the SRA Achievement Tests. The scores obtained were analyzed by means of t-tests. The results of the analyses showed no differences at the .05 level of significance for any of the dependent variables. Drawing a positive conclusion from their findings the authors suggest that generative transform grammar may be as good as traditional grammar in the sixth grade. They offer some recommendations for future research.

8. *College Reading Improvement.* Stebens and Belden (45) attempted to discover how much gain in reading skills was retained from a college reading training program after 5 semesters. The retained gains of the students who completed the reading improvement course were compared with the improved reading performance of students who were enrolled in the college the same 5 semesters, but who did not take the course. The 108 Ss in each group were equated on the basis of scholastic ability as measured by the ACT, initial reading skills as measured by the Nelson-Denny Reading Test, and sex. T-tests were employed to determine any significant differences on the Nelson-Denny Reading Test between (1) pre- and post-training scores for students who took the course, (2) post-training and retest scores for students 5 semesters after taking the course, (3) pretest and retest scores for students who did not take the course, and (4) the gains recorded for pretest and retest scores after 5 semesters for students who did and students who did not take the course. Significant improvements were found for both groups on vocabulary and rate scores, but not for comprehension. Although both groups realized significant improvement in reading skills (except comprehension) after 5 semesters, the group that completed the reading course made comparatively greater gains. The authors conclude that although reading appears to improve after 5 semesters of college work without training, formal training appears significantly to affect reading ability. The authors do not describe the reading program the experimental group received.

In a 2-year study Bahe (2) investigated the effect of reading study skills instruction on the scholastic performance of college underachievers. The Ss of the study had graduated in the upper 30 percent of their high school classes and had attained a composite ACT score of at least 23. However, they received a grade point average of below C during their first college semester. The first year of the experiment the investigator studied thirty-three volunteers from 123 eligible freshmen. The second year the investigator studied twenty volunteers from 118 who were eligible. The author points out that the shortage of volunteers prevented the establishment of desired controls. At the end of 8-week summer reading improvement courses, retesting with the Iowa Silent Reading Tests indicated many initial mean gains. However, follow-up testing found these gains to be temporary for all reading skills except rate, which according to t-test analyses retained a significant improvement. The author concludes that for reading improvement courses to be of lasting value to college under-achievers, the work of the reading-study skills teacher must be systematically reinforced by content area teachers.

Dubois (13) looked at the relative effects of college reading programs that relied upon general developmental reading material, and those that used subject matter material. Fifty students enrolled in a college reading improvement course were randomly assigned to each experimental group. Sixty students who were eligible to take the course, but who elected not to, were identified as a control group. Growth was measured according to the Ss' responses to cloze tests constructed from subject matter textbooks, the Diagnostic Reading Test, and the Iowa Silent Reading Test. A one-way analysis of variance was employed to determine any significant differences among the three groups. A significant difference at the .01 level of confidence was found. The Scheffé method of comparison was used to find whether statistically significant differences occurred in mean gains between the pretest and posttest scores on the three evaluation measures. Analysis of the data revealed that students in both experimental groups made significantly greater gains than the students in the control group. The differences between the two experimental groups were not significantly different at the .05 level of confidence. The authors concluded that reading improvement courses are valuable for college students and that general reading materials and subject-matter type materials were equally effective for improving textbook comprehension and general story or work-study type skills as they were measured in this study.

Gilman and Moreau (19), using thirty-six community college students with low academic ability, investigated the effects of reducing the verbal content in a Computer-Assisted Information (CAI) program on (1) instructional time, (2) comprehension, and (3) the correlation between posttest performance and verbal intelligence. To lower the verbal content, the program used was rewritten in short, concise sentences, and the unusual words were replaced with more common words. The Ss were randomly assigned to two treatment groups. On a posttest of eighteen items constructed by the investigators there was a slight, but not statistically significant difference in favor of the students in the low verbal content group. The time required for instructional purposes was lower at the .05 level of significance for the students who used the program with lowered verbal content. The correlation between posttest score and the California Mental Maturity Test score for verbal intelligence was not significant for the low verbal content group, but was significant at the .05 level of significance for the high verbal content group. On the basis of their findings, the authors recommend greater efforts to program materials with as low verbal content as possible for students with generally poor verbal ability.

Mathis, Smith, and Hansen (31) assessed the attitudes of college students toward CAI before and after they experienced it. Forty-seven females and seventeen males were randomly assigned within sex to two experimental and two control groups. Students in the experimental groups received CAI, and students in the control groups read the same information. One experimental and one control group were pretested on a 30-item attitude inventory constructed to measure attitudes toward CAI. Each of the four groups was further divided into subgroups A and B. Students in the A subgroups received either CAI or reading material which was discussed in class and on which they were to be tested that week. Students in B subgroups received CAI or reading material that had not been covered in class and on which they would not be tested for several weeks. All students were tested immediately after their experience with the attitude inventory. On the basis of the findings of a number of statistical analyses that were performed the authors concluded that:

College students generally have positive attitudes to computers and exposure to CAI in this study increased their positiveness. But the magnitude of the attitude change was dependent upon the kind of experience they had. Those students who had the familiar and relevant CAI program, and made few errors while taking it, showed the greatest increase in positiveness to CAI.

Charles (11) reported the results of a survey of reading service being offered in the community college in the State University of New York. Eighteen of the twenty-one colleges that responded indicated they were providing reading services. The different services being offered were remedial reading, developmental reading, study skills training, and speed reading. Only two colleges reported that their courses in reading gave academic credit.

D. Assessment and Prediction of Reading Achievement

1. *Prediction of Reading Achievement.* Personke and Davis (37) administered the Metropolitan Readiness Tests in English and in Spanish to thirty-eight Spanish-speaking children who were beginning their first year of school. In May of the same school year, the Ss were administered the Metropolitan Achievement Tests, Primary I Battery, in English. After a careful analysis of the correlations between the Metropolitan subtests on Word Knowledge, Word Discrimination, and Reading and the six subtest scores and the total score for both the English and Spanish administrations of the Metropolitan Readiness Tests, the investigators concluded that administrations of the Metropolitan Readiness Tests in English and in Spanish seemed to yield similar predictive relationships with first-grade reading achievement.

Olson and Johnson (36) undertook an investigation to determine if the Frostig Developmental Test of Visual Perception (FDTVP) measures several different perceptual abilities, and to determine how well the instrument predicts reading achievement from first to third grade. To accomplish these objectives, data were collected on 218 beginning first-graders with the following tests: (a) FDTVP; (b) Gates Reading Readiness Test (GRRT); (c) Metropolitan Readiness Test (MRT); (d) Olson Reading Readiness Test (ORRT). In addition, the Ss took the Wechsler Intelligence Scale for Children (WISC) in December of first grade, the Stanford Achievement Test in May of first grade, and the Stanford Word Reading and Stanford Paragraph Meaning tests in May of third grade. To determine the construct validity of the FDTVP, a factor analysis of its five subtests along with the subtests of the other readiness and intelligence measures was undertaken. The results of the oblique factor structure indicated that the FDTVP was a unifactor test. To test the predictive validity of the FDTVP for the first- and third-grade reading achievement, a stepwise regression analysis was employed and the FDTVP was found to produce multiple correlation coefficients of .57 with first-grade word reading, .55 with third-grade word reading, .55 with first-grade paragraph meaning, .58 with third-grade paragraph reading. Since ORRT, GRRT, and a combination of the MRT and WISC produced somewhat higher multiple R's with the criteria, Olson and Johnson concluded that FDTVP was the least useful predictor of reading achievement of those studied.

2. *Test Norms.* Briggs and Tellegen (8) reported scores on the eye-hand coordination subtest of the FDTVP for 360 males and females from ages 3 to 39. Their results agreed with the published norms for ages 4 through 7; however, they found differences between their results and the published norms after age 7. The published norms, according to Briggs and Tellegen, rise too rapidly starting at age 7 and level off too early at approximately age 9. On the basis of a series of studies comparing the reading achievement of children of infant school age (CA's of approximately 5–8 years) on tests with norms of various vintages, Bookbinder (7) developed a set of adjusted norms for the Schonell Graded Word Test up to a reading age of 7½. Although the investigator made no claim for the complete accuracy of the modified norms, he observed that the modified norms were more realistic than the norms presently used with the Schonell Test.

3. *Cloze Tests.* In order to explore the relationship of flexibility of closure and speed of closure to different types of cloze tests, Ohnmacht, Weaver, and Kohler (35) secured a sample of 113 high school students. Subjects were administered four cloze tests representing structural, lexical, abstract noun, and concrete noun deletion patterns as well as measures of flexibility of closure, speed of closure, associational fluency, and verbal comprehension. Correlation and factor analysis techniques were applied to the data. In general, the investigators found that all the analyses showed some relationship of cloze factors to classical closure factors—flexibility of closure and speed of closure, with flexibility more closely related than speed. The relationship was not strong enough, however, to lead the investigators to accept the classical gestalt explanation of visual closure phenomena as descriptive of the process of completing cloze blanks. In addition, it was found that the cloze tests could be differentiated into at least two discriminable dimensions. The lexical deletion, abstract noun deletions, related to vocabulary, was one factor, while the structural, concrete noun deletions were together in another factor.

In a replication of previous studies by Bormuth, Rankin, and Culhane (38) compared the cloze and multiple-choice test scores for 105 fifth-grade pupils. Data were collected during two class periods 4 days apart. On the first day, Ss took one of five cloze tests which were constructed by deleting every fifth word in a selection. The second day they read the unmutilated 275-word passage upon which the cloze test had been based and took a multiple-choice test covering the passage. The cloze test results were correlated with the multiple-choice test result for each passage. Regression coefficients were then computed on each article and multiple-choice percentage scores were predicted from cloze test percentage scores at every 5 percentage points from 50 percent through 100 percent. Then, the cloze test scores predictive of the specified multiple-choice scores were averaged across all five passages. The results can be summarized as follows: (a) correlations between the five cloze and multiple-choice tests ranged from .54 to .77; (b) cloze percentage scores of 41 and 61 predicted multiple-choice percentage scores of 75 and 90, respectively; (c) the latter finding was reasonably consistent with Bormuth's 1968 results, for he found that cloze percentage scores of 44 and 57 predicted the two criteria. Although the investigators concluded that they had cross-validated Bormuth's work, a more conservative conclusion might have been in order.

McLeod and Anderson (32) reported a series of experiments which they felt consistently indicated that certain parameters based on the concepts of empirical estimation of redundancy of printed English could be derived from cloze-type tests. As a result, they proposed that through the use of an Uncertainty Reduction Index (URI) to score cloze tests rather than the use of the proportion of correct restorations, a reliable, valid expression of reading ability can be obtained. The use of URI rather than orthodox cloze scores means that equal weight is not given to all items. URI responses to items, where the correct word is ambiguous, carry little weight, whereas items which are unambiguous carry greater weight. Curiously, the investigators close their report by rhetorically wondering if the extra effort to compute URI is worthwhile, since it shows virtually perfect correlation with the usual cloze raw score. The answer for them is obviously yes, since their last sentence suggests that perhaps the correlation between URI and raw scores should be interpreted as an overdue validation of raw scores.

4. *Test Development and Analysis.* In an attempt to identify unique dimensions of reading comprehension, Schreiner, Hieronymus, and Forsyth (42) constructed an experimental reading test battery with eight subtests: (a) Speed of Noting Details, (b) Speed of Reading, (c) Paragraph Meaning, (d) Determining Cause and Effect, (e) Reading for Inference, (f) Selecting Main Ideas, (g) Verbal Reasoning, and (h) Listening Comprehension. The subtests were organized into two half-length forms and were administered 1 week apart to 513 fifth-grade pupils. In addition, the Vocabulary and Reading Tests from the Iowa Tests of Basic Skills (ITBS) were included in the analysis, since the Ss had taken them a week before the study began. Once the data were collected, subtest intercorrelations were run and corrected for attenuation, and the reliabilities of differences between all possible pairs of subtests were calculated. The results of both analyses were similar: (a) subtests of the experimental battery requiring speed, listening, or verbal reasoning ability were relatively unique when compared with each other and when compared with the other experimental subtests; (b) subtests labeled Main Ideas, Paragraph Meaning, Cause and Effect, and Inferences were shown to be highly interrelated but relatively unique when compared with the other experimental subtests; (c) ITBS subtests were highly intercorrelated but relatively independent statistically from subtests in the other categories. Thus, the investigators concluded that of the subtests studied only the Speed of Reading, Listening Comprehension, Verbal Reasoning, and Speed of Noting Details subtests were reasonably independent enough to be used diagnostically.

Weaver, Bickley, and Ford (49) compared standard reading comprehension test items that were dependent upon reading a passage prior to their successful completion with items that had a very loose relationship to the accompanying reading paragraph and that could be answered without reading the paragraph. Data were collected on 148 college Ss who were randomly assigned to one of four, 40-item tests: (a) Reading paragraph present; more paragraph relatedness among the items; (b) Reading paragraph present; less paragraph relatedness among items; (c) Reading paragraph absent; more paragraph relatedness among the items; (d) Reading paragraph absent; less paragraph relatedness among the items. A two-way analysis of variance with fixed factors indicated significant differences ($p < .05$) for the main effects of Item Dependency and Paragraph Condition. In essence, the investigators found that mean achievement on their tests was higher when the paragraphs were present and when the test items were less dependent on the content of the paragraphs.

Berger (6) reported the following reliability information for the rate and comprehension portions of the Braam and Sheldon Flexibility of Reading Test: (a) an overall rate reliability of .89 between Forms 1 and 2; (b) an overall rate reliability of .90 between Forms 2 and 3; (c) an overall comprehension reliability of .66 between Forms 1 and 2; (d) an overall comprehension reliability of .52 between Forms 2 and 3. The rate and comprehension reliability coefficients were based on undescribed samples of seventy and sixty-three college freshmen, respectively. The investigator drew no conclusions from his findings.

E. Remedial and Corrective Reading Instruction

Carlin (10) assessed the effects of an intensive program of clinical reading instruction given for one month in a university environment to five sixth-grade boys who possessed severe reading problems. At the beginning of the study, each boy was randomly assigned to a team of five reading specialists. Each of the teams devoted their full attention to planning and conducting an instructional program designed to meet the needs of its boy. By comparing results of final evaluations with those made prior to the one month experiment, Carlin found that general progress was made with each S in the areas of word analysis, vocabulary, study techniques, and comprehension. In addition, he noted that monthly mean gains in general reading ability, as measured by the California Achievement Test, were above expected performance for each S for the year in which the experiment took place.

To determine the effects of the Delacato recommended training on the reading achievement and visual-motor integration of disabled readers with crossed or uncertain lateral expression, O'Donnell and Eisenson (34) randomly selected sixty children with these characteristics and assigned them to one of three training groups, which met for 30 minutes daily for a period of 20 weeks. Group I participated in cross-pattern creeping for 10 minutes, cross-pattern walking for 10 minutes, visual pursuit for 2 minutes, and filtered reading for 5 minutes during each 30 minute session. Group II participated in visual pursuit activity for 2 minutes, and selected physical education activity for 15 minutes during each training session. Group III engaged in selected physical education activity during the entire training period. Mean gains between pretests and posttests on the Gray Oral Reading Test, the Developmental Test of Visual-Motor Integration, and the seven subtests of the Stanford Diagnostic Reading Test constituted the dependent variables. In general, the investigators found that the Ss who were given either the Delacato recommended training, or the modification of this training did not make significantly greater gains in reading achievement or visual-motor integration than did pupils who were limited to the physical education program.

Frager and Stern (18) studied two types of tutor-instruction to determine which approach would produce the most growth in tutors as well as tutees. The tutors in this instance were forty-eight sixth-grade students, divided into two equal groups of high and low achievers on the basis of reading scores on the Stanford Achievement Test. Each tutor was paired with a kindergarten child who had been tested and found to be in need of remedial work; however, before tutoring, the tutors were randomly assigned to one of two treatments designed to help them work with young children. Treatment I consisted of traditional instructional procedures in which the tutorial process was described, suggestions for working with the younger children were presented, and questions on specific problems were answered. In Treatment II, tutors were taught a procedure which consisted of five basic steps: defining goals, defining obstacles, specifying alternatives, identifying consequences of specific alternatives, and making selections among alternatives. Using the criterion test provided with the McNeil ABC Learning Activities as a pre- and post-measure, Frager and Stern found that the kindergarten children who received tutoring, whether from Treatment I or II tutors, were superior to the children who did not receive tutoring. In addition, the experience was very worthwhile for the tutors, particularly in terms of their morale, attitudes, school attendance, and their feelings about themselves.

In an effort to assess the relative effects of group counseling on reading achievement, Thayer (47) assigned twenty-one eighth-grade boys, with normal IQ scores and low reading achievement, to three treatment groups. The seven Ss in Group A received group counseling once a week for 45 minutes; in addition, the mothers of the Ss in group A met in separate weekly group sessions. The seven Ss in Group Z participated in weekly group counseling sessions, but their mothers did not. The remaining Ss were assigned to Group C, which did not meet. After 10 weeks, gain scores on alternate forms of the Gates Reading Survey, which were administered before and after the experiment, were compared across groups. It appeared that Group A had the largest mean increase, Group Z had the second largest increase, and Group C had the smallest increase. Unfortunately, no tests of significance were applied to the data and the investigator could simply conjecture as to their meaning.

Graeme and Harris (21) found that training in the use of dictionary pronunciations and practice in oral reading can significantly increase both the number of words attempted by retarded readers and the number they could pronounce correctly.

REFERENCES ON THE TEACHING OF READING

1. **Affleck, Muriel A.**, "Phonic Principles in Intermediate Reading," *Reading Improvement* 6:(no. 3)69–77, Winter 1969, IV.B.
2. **Bahe, Virginia R.**, "Reading—Study Instruction and College Achievement," *Reading Improvement,* 6:(no. 3)57–61, Winter 1969, IV.C.8.
3. **Bartolome, Paz I.**, "Teachers' Objectives and Questions in Primary Reading," *The Reading Teacher,* 23:(no. 1)27–33, October 1969, IV.A.
4. **Beck, Ray; Talkington, Larry W.**, "Frostig Training With Headstart Children," *Perceptual and Motor Skills,* 30:521–522, 1970, IV.C.1.
5. **Belcastro, Frank P.**, "Programmed Algebra and Reading Achievement," *The High School Journal,* 53:(no. 5)310–317, February 1970, IV.C.7.
6. **Berger, Allen**, "The Reliability of the Flexibility of Reading Test," *The Alberta Journal of Educational Research,* 15:(no. 3)165–168, September 1969, IV.D.4.
7. **Bookbinder, G. E.**, "Variations in Reading Test Norms," *Educational Research,* 12:(no. 2)99–105, February 1970, IV.D.2.
8. **Briggs, Peter F.; Tellegen, Auke**, "Further Normative Data on a Frostig Subtest, Eye-Hand Coordination," *Perceptual and Motor Skills,* 30:640–642, 1970, IV.D.2.
9. **Brody, Erness Bright**, "Achievement of First- and Second-Year Pupils in Graded and Nongraded Classrooms," *The Elementary School Journal,* 70:(no. 7) 391–394, April 1970, IV.C.4.
10. **Carlin, James B.**, "Intensive Reading Instruction," *The Reading Teacher,* 23:(no. 5)422–425, February 1970, IV.E.
11. **Charles, Harvey**, "A Survey of Reading Services in the Community Colleges in the State University of New York," *Journal of the Reading Specialist,* 9:(no. 3)111–114, March 1970, IV.C.8.
12. **Downing, J.; Latham, W.**, "A Follow-Up of Children in the First i.t.a. Experiment," *The British Journal of Educational Psychology,* 39:303–305, 1969, IV.C.6.
13. **Dubois, Ronald L.**, "Improvement of Textbook Comprehension in College Reading Classes," *Journal of Reading,* 13:(no. 2) 113–118, 165, November 1969, IV.C.8.
14. **Durkin, Dolores**, "A Language Arts Program for Pre-First-Grade Children: Two-Year Achievement Report," *Reading Research Quarterly,* 4:534–565, Summer 1970, IV.C.2.
15. **Falik, Louis H.**, "The Effects of Special Perceptual-Motor Training in Kindergarten on Reading Readiness and on Second Grade Reading Performance," *Journal of Learning Disabilities,* 2:(no. 8)395–402, August 1969, IV.C.1.
16. **Farr, Roger; Laffey, James; Brown, Rexel**, "Secondary Reading Programs in Indiana: Status and Needs," *Journal of Reading,* 13:269–274, 317, January 1970, IV.A.
17. **Feldhusen, Hazel J.; Lamb, Pose; Feldhusen, John**, "Prediction of Reading Achievement Under Programmed and Traditional Instruction," *The Reading Teacher,* 23:(no. 5)446–453, February 1970, IV.C.3.
18. **Frager, Stanley; Stern, Carolyn**, "Learning by Teaching," *The Reading Teacher,* 23:(no. 5)403–405, 417, February 1970, IV.E.
19. **Gilman, David Alan; Moreau, Nancy Ann**, "Effects of Reducing Verbal Content in Computer-Assisted Instruction," *AV Communication Review,* 17:(no. 3) 291–297, Fall 1969, IV.C.8.
20. **Goudey, Charles E.**, "Reading—Directed or Not?" *The Elementary School Journal,* 70:245–247, February 1970, IV.C.5.
21. **Graeme, Sharron C.; Harris, Mary B.**, "Improving Word Recognition in Retarded Readers," *The Reading Teacher,* 23:(no. 5) 418–421, February 1970, IV.E.
22. **Jensen, Norma J.; King, Ethel M.**, "Effects of Different Kinds of Visual-Motor Discrimination Training on Learning to Read Words," *Journal of Educational Psychology,* 61:(no. 2) 90–96, 1970, IV.C.2.
23. **Kasdon, Lawrence M.; Kelly, Dean**, "Simulation: In-Service Education for Teachers of Reading," *The Journal of Experimental Education,* 38:(no. 1) 79–86, Fall 1969, IV.A.
24. **Keim, Richard P.**, "Visual-Motor Training, Readiness, and Intelligence of Kindergarten Children," *Journal of Learning Disabilities,* 3:(no. 5)19–22, May 1970, IV.C.1.
25. **Keipe, Ashtoreth, Wood, Roger**, "A Comparative Study of Achievement Between a Linguistics Program of Generative Transform Grammar and a Traditional Program of Grammar," *Elementary English,* 47:(no. 4)535–539, April 1970, IV.C.7.
26. **Kingston, Albert J.; Cowart, Harry J.**, "Teacher Use of Reading in Preparing High School Study Assignments: 1919 and 1969," *Journal of Reading Behavior,* 1:(no. 4)3–9, Fall 1969, IV.A.
27. **Kresh, Esther**, "Comparative Effectiveness of Initial Teaching Alphabet versus Traditional Orthography in First Grade in the Pittsburgh Public Schools," *The Journal of Educational Research,* 63:(no. 4) 170–172, December 1969, IV.C.6.
28. **LaConte, Christine**, "Reading in Kindergarten," *The Reading Teacher,* 23:(no. 2)116–120, November 1969, IV.A.
29. **Lam, Charlotte Dawson**, "Pupil Preference for Four Art Styles Used in Primary Reading Textbooks," *The Reading Teacher,* 23:((no. 2)137–143, November 1969, IV.B.
30. **Lockmiller, Pauline; DiNello, Mario C.**, "Worlds in Color Versus a Basal Reader Program With Retarded Readers in Grade 2," *The Journal of Educational Research,* 63:(no. 7)330–334, March 1970, IV.B.
31. **Mathis, Arthur; Smith, Timothy; Hansen, Duncan**, "College Students' Attitudes Toward Computer-Assisted Instruction," *Journal of Educational Psychology,* 61:(no. 1)46–51, 1970, IV.C.8.
32. **McLeod, John; Anderson, Jonathan**, "An Approach to Assessment of Reading Ability Through Information Transmission," *Journal of Reading,* 2:(no. 2) 116–141, Spring 1970, IV.D.3.
33. **Meisel, Stephen; Glass, Gerald G.**, "Voluntary Reading Interests and the Interest Content of Basal Readers," *The Reading Teacher,* 23:(no. 7)655–659, April 1970, IV.B.
34. **O'Donnell, Patrick A.; Eisenson, Jon**, "Delacato Training for Reading Achievement and Visual-Motor Integration," *Journal of Learning Disabilities,* 2:(no. 9)441–446, September 1969, IV.E.
35. **Ohnmacht, Fred W.; Weaver, Wendell W.; Kohler, Emmett T.** "Cloze and Closure: A Factorial Study," *The Journal of Psychology,* 74:205–217, 1970, IV.D.3.
36. **Olson, Arthur V.; Johnson, Clifford I.**, "Structure and Predictive Validity of the Frostig Developmental Test of Visual Perception in Grades One and Three," *The Journal of Special Education,* 4:(no. 1)49–52, Winter-Spring 1970, IV.D.1.

37. **Personke, C. R.; Davis O. L.,** "Predictive Validity of English and Spanish Versions of a Readiness Test," *The Elementary School Journal*, 2:79–85, November 1969, IV.D.1.
38. **Rankin, E. F.; Culhane, J. W.,** "Comparable Cloze and Multiple-Choice Comprehension Test Scores," *Journal of Reading*, 13:193–198, December 1969, IV.D.3.
39. **Reichard, Cary L.; Reid, William R.,** "An Investigation of Format for Reading Material for the Educable Mentally Retarded," *Journal of Reading*, 13:(no. 5)363–366, February 1970, IV.B.
40. **Rystrom, Richard,** "Dialect Training and Reading: A Further Look," *Reading Research Quarterly*, 5:(no. 4)581–599, Summer 1970, IV.C.3.
41. **Schnayer, Sidney W.; Robinson, Leona A.,** "An Analysis of Phonic Systems for the Primary Grades in Eight Basal Reader Series," *Journal of the Reading Specialists*, 9:(no. 2)58–72, December 1969, IV.B.
42. **Schreiner, Robert L.; Hieronymus, A. N.; Forsyth, Robert,** "Differential Measurement of Reading Abilities at the Elementary School Level," *Reading Research Quarterly*, 1:84–99, Fall 1969, IV.D.4.
43. **Smith, Richard J.; Burns, Thomas,** "The Effects of Different Instructional Practices on Student Enjoyment and Interpretation of a Ballad," *Journal of Reading*, 13:(no. 5)345–354, February 1970, IV.C.7.
44. **Smith, Richard J.; Bragstad, Bernice; Hesse, Karl D.,** "Teaching Reading in the Content Areas—An In-service Model," *Journal of Reading*, 13:(no. 6)421–428, March 1970, IV.A.
45. **Stebens, Duane L.; Belden, Bernard R.,** "Retention of Gains in Reading After Five Semesters," *Journal of Reading*, 13:(no. 5)339–343, February 1970, IV.C.8.
46. **Stefflre. Buford,** "Run Mama, Run: Women Workers in Elementary Readers," *Vocational Guidance Quarterly*, 18:(no. 2)99–102, December 1969, IV.B.
47. **Thayer, J. A.,** "Johnny Could Read—What Happened?,"*Journal of Reading*, 13:(no. 7)501–506, 561, April 1970, IV.E.
48. **Warner, Dolores,** "The Role of Pupil Judgment in Reading Instruction," *The Reading Teacher*, 23:(no. 2)108–111, November 1969, IV.C.3.
49. **Weaver, W. Wendall; Bickley, A. C.; Ford, Fraughton, G.,** "A Cross Validation Study of the Relationship of Reading Test Items to Their Relevant Paragraphs," *Perceptual and Motor Skills*, 29:11–13, 1969, IV.D.4.
50. **Williams, Joanna P.,** "Training Kindergarten Children to Discriminate Letter-Like Forms," *American Educational Research Journal*, 6:(no. 4)501–513, November 1969, IV.C.1.

Forthcoming Articles

Bigelow, Gordon S.; Kent State University, Kent, Ohio
"FIELD DEPENDENCE–FIELD INDEPENDENCE IN 5- TO 10-YEAR-OLD CHILDREN"

Booney, Merl E.; North Texas State University, Denton
"ASSESSMENT OF EFFORTS TO AID SOCIALLY ISOLATED ELEMENTARY SCHOOL PUPILS"

Grant, Claude W.; University of Utah, Salt Lake City
"FACULTY ALLOCATION OF EFFORT AND STUDENT COURSE EVALUATIONS"

Keefer, Karl E.; University of Tennessee at Martin
"CHARACTERISTICS OF STUDENTS WHO MAKE ACCURATE AND INACCURATE SELF-PREDICTIONS OF COLLEGE ACHIEVEMENT"

Kennedy, Thomas G. and Robert A. Humphrey; Western Montana College
"THE EFFECT OF THE SYSTEM APPROACH ON CHANGING PRE-SERVICE TEACHER ATTITUDES TOWARD SELECTED INSTRUCTIONAL DESIGN FACTORS"

Nelson, Dennis Earl; American Institutes for Research, Palo Alto, California
"THE COLLEGE ENVIRONMENT: ITS MEANING TO ACADEMICALLY SUCCESSFUL AND UNSUCCESSFUL UNDERGRADUATES"

Null, Eldon J.; Purdue University, Lafayette, Indiana
"RELATIONSHIPS BETWEEN PERSONAL VARIABLES OF TEACHERS AND THEIR PERCEPTION OF THE BEHAVIOR OF SCHOOL PERSONNEL"

Raina, M. K.; University College of Education, Indore, India and Raina, T. N.; Regional College of Education, Ajmer, India
"PERCEPTION OF TEACHER–EDUCATORS IN INDIA ABOUT THE IDEAL PUPIL"

Sebok, Richard F.; University of Texas at El Paso
"THE EFFECT OF SELECTED DEMOGRAPHIC FACTORS ON PREDICTION OF GRADE–POINT AVERAGE BY MALE FRESHMAN STUDENTS"

Tosi, Donald J., Richard D. Briggs, and Rosemary M.McCaulley; Western Michigan University, Kalamazoo
"STUDY HABIT MODIFICATION AND ITS EFFECT ON ACADEMIC PERFORMANCE: A BEHAVIORAL APPROACH"

Wright, E. N. and R. S. Rogers; Board of Education for the City of Toronto, Canada
"A STUDY OF CHILDREN'S DRAWINGS OF THEIR CLASSROOMS"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 6, February 1971)

Intelligent Word Associations in High School Students

PHILIP LAMBERT, LEE H. HANSEN, and EDGAR F. BORGATTA
The University of Wisconsin, Madison

ABSTRACT

The place of word associations in measures of verbal abilities is discussed. A study involving a sample of high school students selected to be representative of a broad and "normal" population in abilities is reported. The test devised of word associations is found to be reliable in this sample, and is found not only to be well related to measures of verbal abilities, but equal to them in ability to predict the external criterion of level of English class to which the student has been assigned.

IN A RECENT paper Borgatta (1) reported the serendipitous finding that a word association test (WAF) would measure intelligence reliably. The original study was devised to develop a set of projective tests that would measure personality characteristics, but the method was unable to locate sufficient word association items that would replicate in three samples to warrant development of objective-projective personality tests. The conclusion was that either the relationship between word associations and personality characteristics is insufficiently strong to permit a test to be developed, or the procedures involved in the test development were insufficient at some point, such as having a poor set of criteria, an insufficient array of word associations, or insufficient variation in the populations. However, the procedures were deliberately designed to satisfy many of the prior criticisms in the development of personality tests based on objective-projective measures, that is, when the projective personality test would depend strictly on the existence of correlates to personality variables in the population.

In the design of this study, one major concern was in the possible effects of verbal ability that could interfere with the discovery of word associations that correlated with personality variables. It is through this interest that the relationships between the word associations and verbal ability were discovered. The development of the test of intelligent word associations was unexpected since it occurred with a prior tradition of failures to locate correlates to verbal ability in the study of word association structures (4). It also was surprising since the source of the items was from "misleaders" in verbal abilities tests. The "misleaders" are generally used as the *incorrect* answers, that is, as associations instead of definitions, in the normal task given in the design of verbal ability which requires the production of synonyms or definitions of concepts.

Some Theoretical Issues

In the original report, the basis for the development of the word associations test was in three relatively homogeneous samples of different ability levels. A major question left unresolved after the initial development of the test was its generality in the population. This is not merely a question of external validity, but one that raises the issue of whether or not, in the general population, associations are an important part of the processes generally subsumed as verbal ability. While the word associations were located by an arbitrary procedure of correlation analyses—independently from analysis of the content of the words—after the word associations were examined, it became clear that "correct" word associations did indeed make sense. The nature of the associations were that two words could be tied together by a plausible story or "story-line." So, for example, there may be an intelligent association between *swig* and *grog*, as: "One swigs grog." Examination of the array of word associations indicates that the connection between two words by a story-line seems to be the way that the "correct" answers are found, although in some cases this may be more accurately defined as saying that the incorrect answer arises because for some word pairs, no story-line can be found. The question thus is raised as to the nature of the associations themselves, and possibly definitions (synonyms) should be considered as one form of associations, the form which is given priority in communicating the content of the language. So, a more direct association for *swig* could be *tope*, but to make such distinctions, one would indeed have to know the language, and it easily is seen how a more difficult "correct" item could be built that involves both a synonym and a "correct" story-line association. The notion of verbal ability that arises is one defined by the amount of knowledge one has in the language that permits definition of concepts, where definition is a relative term ranging from synonym to associations. Presumably, associations could be viewed as prior to the ability to provide synonyms. There is an analogy in science that sometimes people "talk around" a topic before they are able to address it directly.

The Current Study

The current study was designed to provide a sample base having several characteristics. First, the ability level selected was to include a range such as might be expected in the general population. The means for achieving this was to select a high school in which the ability range was judged by school authorities to include a broad array of abilities generally within the normal performance levels that might be expected in a Midwest city. The data collection permitted building-in certain other variables of interest. Data were collected through English classes where students were grouped on prior performance (high, middle, and low), and each of the 4 high school years provided such classes. Thus, the sample had limited age variability, and a global judgment on knowledge of English was available as a measure of external validity. Two verbal abilities tests were administered at the same time, the Verbal Power Test (VPT) (3), a word-matching test with 150 items split in two halves, and one-half of the Quick Word Test (QWT) (2). The tests of verbal ability were selected because they are known to be highly reliable and can be administered in a very short time period. All three tests were administered in one class period, and all forms returned to the researchers were judged to be scorable.

Table 1.—Intercorrelations Among Verbal Abilities, Word Associations Form, Age, and Level of English Class Assignment (N=269)

	VPT A	VPT B	VPT Total	QWT	WAF	Age	English Level
VPT A Half	—	.80	.xx	.70	.68	.12	.63
VPT B Half		—	.xx	.67	.67	.02	.58
VPT Total			—	.72	.71	.07	.63
QWT (Half)				—	.55	-.05	.60
WAF					—	.18	.62
Age						—	-.02
English Level							—

The mean score for the VPT was 91.0 correct with a standard deviation of 16.5, and mean score for the QWT was 32.8 with a standard deviation of 7.5. The characteristics of the sample on these bases are similar to those that might be projected for the population of the United States on the basis of the information provided in the test manuals.

Table 1 presents the intercorrelations among the verbal abilities tests, the WAF score, and age and level of English class assignment. The intercorrelation between the halves of the VPT is indicated to be .80. Thus, the reliability of the total form can be expected to be higher for this sample. The correlation between the VPT total score and the half of the QWT used for this study is .72. The correlation between the VPT and the word associations score is .71, approximately at the same level as the correlation between the two verbal abilities scores. The correlation between the word association score and the half of the QWT score is .55, suggesting that these scores are either less reliable or less related to each other than to the VPT score. Since the reliability of the QWT has been reported as extremely high, the hypothesis of lower reliability is less likely, and probably the WAF has more in common with the task of the VPT than with that of the QWT. However, independently of this judgment, what is of interest is the high level of correlation for the WAF generally, which suggests that its reliability must be quite high. The Cronbach Alpha computed for the word associations score for this sample was found to be .75, which is higher than the previously reported values for the samples on which the test was developed. The several verbal tests do not correlate substantially with age for this group (the standard deviation for age was 1.5 years), but the correlation to the level of English class assignment is quite substantial. Examining the last column of Table 1, it is clear that the verbal abilities tests and the word associations score are all correlating with this more global external judgment of ability in English at about the same level.

REFERENCES

1. Borgatta, E. F., "Intelligent Word Associations," University of Wisconsin, Madison, February 1970, mimeo.
2. Borgatta, E. F.; Corsini, R. J., *Quick Word Test: Manual*, Harcourt, Brace, and World, New York, 1964.
3. Francesco, E., *The Verbal Power Test of Concept Equivalents, VPT Manual*, Western Psychological Services, Beverly Hills, California, 1963.
4. Moran, L. J.; Veldman, D., "Relationship of Association Structures to Verbal Ability," *Psychological Reports*, 20:839–842, 1967.

Book Reviews

ROBERT E. CLASEN
Book Review Editor

Stauffer on Reading: Two for the Show

Directing Reading Maturity as a Cognitive Process

Stauffer, R. G. Harper and Row, New York, 1969, 498 pp. $11.75.

THIS BOOK is clearly not just another reading methods text. Prospective readers who are looking for a straightforward guide to the teaching of reading are well advised to look elsewhere. On the other hand, prospective readers who are looking for a carefully developed, albeit sometimes ponderously presented, philosophy of reading instruction need look no further. The author's stated purpose is to "*. . . acquaint graduate students with a philosophy of reading instruction.*" Acquaint them he does.

Stauffer's basic premise, which is shared by virtually every other worker in the field, is that reading must be viewed as a thinking process. His essential corollary, which is not so universally shared in practice, is that it must be taught as a thinking process at all levels of instruction. The vehicle suggested for pursuing these ends is the Directed Reading-Thinking Activity (DR-TA), an extension of the Directed Reading Activity originally suggested by Betts, which comprises the following elements in a group situation: (1) identification of pupil purposes for reading; (2) adjustment of rate to purpose and to the nature and difficulty of material; (3) teacher observation of pupil performance for the purpose of noting abilities to adjust rate, checking on and providing help with comprehension, and providing immediate help in the use of varied word attack clues; (4) development of comprehension through clarification of purpose, use of additional materials, and/or development of concepts; and (5) skill training activities, e.g., directed attention, reorganization of ideas, vocabulary development, oral reading. The entire book, then, is fundamentally an elaborate rationale for the DR-TA as it is conceived by the author.

While the focus of the book is upon the DR-TA idea, there is no denying that the coverage is comprehensive. This point is best supported by a consideration of the content of the eleven chapters. In Chapter 1, reading is defined as a thinking process and the position that people can be trained to read critically and reflectively is defended. The second and third chapters are devoted to group and individualized approaches to reading instruction; characteristics of each are considered and a complementary arrangement for both is suggested. Beginning reading instruction is considered in Chapters 4 and 5. Readiness is examined in terms of what it is and how teachers can best assist in its development. The language-experience approach is discussed in pragmatic terms; but at times the reader might feel that he is being subjected to an overly long commercial for the Stauffer Word Card Holder, which can be used in the approach. Chapter 6 has to do with the development of word-attack skills; the author makes the point that word-attack skills can be efficiently applied only if the reader is able to grasp the meaning of what is read. Chapter 7 is entitled "Concept Development" and amounts to a paraphrase of a paraphrase of Piaget. The final chapters have to do with the role of libraries in reading instruction, evaluation, the variation of reading rate, and critical and creative reading.

The author himself has stated the main limitation of *Directing Reading Maturity as a Cognitive Process:* ". . . to acquaint graduate students with a philosophy of reading instruction." The reader, then, should not be disappointed when what he finds is indeed a statement of a personal philosophy of reading instruction. Instead, the reader should be pleased that Professor Stauffer has seen fit to expend the time and energy required to develop and record such a statement and to lay it open to careful scrutiny. Having become acquainted with one man's philosophy of reading instruction may be a first step toward developing one's own.

Wayne Otto
The University of Wisconsin—Madison

Teaching Reading as a Thinking Process

Stauffer, Russell G. Harper and Row, New York, 1969. 410 pp. $11.75.

AS STATED in the foreword, the chief concern of the author was to acquaint pre-service and in-service teachers with a philosophy of reading instruction. The book is intended primarily for use as a textbook in college and university undergraduate courses on the foundations of reading instruction. It is good to find an author who keeps reading instruction in perspective by emphasizing the teaching of reading as a thinking process when the current emphasis is upon teaching reading as a decoding process. Although Russell Stauffer recognizes the necessity of teaching letter-sound correspondences, he sees this objective of the reading curriculum as distinctly subordinate to teaching children to respond to ideas encountered in print. Analyzing words, he stresses, is done to discover ideas. Therefore, attention should always be focused on the message of a written communication and not on the words that are used.

Stauffer's recommended instructional approach for teaching reading as a thinking process is to involve students in directed reading thinking activities (DRTA). The essence of this approach is purposeful reading as an outgrowth of pupil interests and experiences. The reader is provided with classroom descriptions of both individual and group instruction based on the DRTA approach.

The book contains eleven chapters, and three of them deserve special praise: "Directed Reading-Thinking Activity Illustrations," "The Language-Experience Approach," and "Developing Skill in Word Recognition." The information in these chapters is presented with clarity and restraint, two qualities often absent in other chapters. Prospective and practicing teachers will undoubtedly find the ideas presented in these three chapters insightful and useful.

(Continued on page 286)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 6, February 1971)

The Schools' Contribution to the Cumulating Deficit[1]

Sunnyuh Shin Chang and James Raths

ABSTRACT

Do school programs contribute to cumulating deficits in academic achievement found between middle-class children and lower-class children through differential emphases given to the basic skills? Fourteen schools, seven serving mostly middle-class children and seven serving mainly lower-class children were identified in a suburban school district near Washington, D. C. An item analysis was performed on the Iowa Tests of Basic Skills battery administered routinely to the fifth grade students in these schools. Items with a positive phi coefficient of .35 or higher were identified as discriminating between the two types of schools. Teachers in the schools were administered a questionnaire asking them to estimate the emphasis their program placed on teaching the basic skills. It was found that middle-class schools placed greater emphasis on the basic skills than did lower-class schools. Also, the middle-class schools placed greater emphasis on the skills that discriminated between the schools than did the lower-class schools, while the emphasis placed on the non-discriminating skills was quite similar.

CHILDREN FROM middle-class homes tend to score higher on standardized achievement tests than those from lower-class homes. This rather reliable finding has been explained by factors of genetics, notions of cultural deprivation, attributions of mental illness, and many other similar constructs. A second rather consistent observation is that the gap between the wealthy and poor children observed on achievement test performances tends to widen with every passing year of school. Attempts to explain this phenomena, termed by many as the cumulating deficit, in terms of out-of-school factors have not been very convincing. The purpose of this study was to identify in-school factors which may account for the variance between the scores of middle-class and lower-class children on standardized achievement tests.

Overview and Hypotheses

Schools serving primarily middle-class children and those serving mainly lower-class students were identified in a Washington, D. C., suburban school district. The Iowa Tests of Basic Skills had been administered to all fifth grade children in the district. An item analysis was performed to determine which test items discriminated between the middle-class and lower-class schools. A random sample of those items that discriminated and those that did not discriminate was selected and entered on a questionnaire administered to fourth grade teachers in both sets of schools. The questionnaire asked teachers to estimate the degree of emphasis to each of the items. It was hypothesized that:

1. *Within the same school system, teachers in the lower-class schools report different curricular emphasis in terms of time spent on items in the skills area than do teachers in the middle-class schools.*
2. *The differences in emphasis reported by teachers in lower-class schools and teachers in middle-class schools are not similarly reflected for the items that discriminate and do not discriminate between the two school groups.*

Definition of Terms

The following definitions were adopted for the purposes of this study:

Lower-class schools are schools with fifth grade students whose average percentile score on the Lorge-Thorndike Intelligence Tests, Form 1, Level B, Nonverbal Battery is 28 or below *and* where more than 70 percent of the students have parents with only a high school education or less.

Middle-class schools are schools with fifth grade students whose grade average percentile score on the Lorge-Thorndike Intelligence Tests, Form 1, Level B, Nonverbal Battery is 68 or above *and* where more than 60 percent of the students have at least one parent who attended college.

Discriminating items are items from the Iowa Tests of Basic Skills with positive phi coefficients of .35 or above resulting from an item analysis.

Non-discriminating items are items with positive phi coefficients below .35.

Curricular emphasis is the reported amount of time spent by teachers on topics reflected in the items sampled from the Iowa Tests of Basic

Skills. Teachers indicated the emphasis each topic received on a 4-point Likert scale ranging from: 1—very little time spent to 4—very much time spent.

Procedures

Identification of Schools. All the fifth graders in a Washington, D. C., suburban county public school system were given a group intelligence test, the Lorge-Thorndike Intelligence Test, Form 1, Level B, Nonverbal, and the standardized achievement test, the Iowa Tests of Basic Skills, Form 1, at the beginning of the 1969 academic year in accordance with County policy. To select the middle-class and the lower-class schools, a request was made by the investigators to a high level administrator at the central office of the County's Board of Education to rank the schools according to the grade average percentile scores of the Lorge-Thorndike Intelligence Tests. Seven schools which had grade average percentile scores of 68 or above and seven other schools which had grade average percentile scores of 28 or below were selected for this study. The educational level of parents of each fifth grade student was determined from the school records, and the percentage of parents who attended college, high school, or less than high school were calculated. In each of the seven schools with grade average percentile scores of 68 or higher, more than 60 percent of the parents (one or both) attended college. More than 70 percent of the parents in each of the seven schools with grade average percentile scores of 28 or below had only a high school level education or less. A total of 1,062 fifth graders, 530 children from middle-class and 532 children from lower-class schools, were involved in this study. Table 1 presents the total number of fifth grade students, grade average percentile scores of the Lorge-Thorndike Intelligence Tests, and the percentage of parents who attended college, high schools, or less than high school from the middle-class and lower-class schools. Although two schools, one middle class and the other lower class, did not permit the examination of records for the information concerning the parents' education, the investigator was assured by the county administrator that the two schools met the criteria stated.

Identification of Discriminating Items. To find the items which discriminated between middle-class and lower-class groups of children, the phi coefficient was computed for each of the 461 items in the Iowa Tests of Basic Skills. The items with phi coefficients of .35 or above were identified as discriminating items and classified according to skill areas indicated in the Teacher's Manual (2). Out of forty-three items in the vocabulary subtest, eight were discriminating; of seventy-four items in the reading subtest, eleven were discriminating. There were no discriminating items among either the forty-three items on the spelling subtest or the forty items on the punctuation subtest. Only one discriminating item was found in the forty items on capitalization, and eight discriminating items in the thirty-two items on usage were located. In the twenty-six items dealing with graphs and tables, two items were discriminating; out of fifty-six items on reference materials, two were discriminating. No discriminating items were found in the thirty-six items on the reading map subtest. In a total of forty-two arithmetic concepts subtest items, fourteen were discriminating, and out of the twenty-nine items on the problem solving subtest only one was found discriminating. The computation of percentage of the discriminating items on each subtest of the Iowa Tests of Basic Skills shows the highest percent, 33.33, in the arithmetic concepts subtest followed by 25 percent in the usage subtest, 18.60 percent in vocabulary, and 14.86 percent in reading comprehension subtest.

Preparation of the Questionnaire. The purpose of the Basic Skills Questionnaire was to assess the amount of emphasis a teacher gives to various topics within his instructional program. The ques-

Table 1.—Number of Students, Grade Average Percentile Scores on the Lorge-Thorndike Intelligence Tests, and Parent Education for Middle-class and Lower-class Schools

School	Number of Students	Grade Average Percentile	Parent Education (in percent) High School or less	Parent Education (in percent) College (one or both parents)
Lower Class				
A	67	12	90	10
B	66	19	97	3
C	84	20	*	*
D	87	21	86	14
E	94	25	71	29
F	81	26	88	12
G	53	28	88	12
TOTAL	532			
Middle Class				
A	95	68	20	80
B	52	68	40	60
C	75	71	*	*
D	100	71	28	72
E	39	74	24	76
F	69	77	26	74
G	100	80	29	71
TOTAL	530			

*Information was not available.

tionnaire contained forty statements corresponding to items in the Iowa Tests of Basic Skills, Form 1. All the items with phi coefficients of .35 or above in the Iowa Tests of Basic Skills were selected as discriminating items and divided into specific skills classifications based on the Teacher's Manual classification. From these classifications, twenty were randomly selected for inclusion in the questionnaire as discriminating statements. Twenty non-discriminating items were also randomly selected from the remainder of the skills classifications and included in questionnaire statements. Eight items with negative phi coefficients ranging from —.009 to —.078 were not included among the non-discriminating items. The questionnaire was given to all forty-four fourth grade teachers in the fourteen schools involved in this study in the middle of May 1970. Each item was rated from one to four, to indicate that the specific skill area presented in the item was (1) emphasized very little, (2) emphasized some, (3) emphasized much, or (4) emphasized very much. Using the Spearman–Brown formula, a reliability of .91 was estimated for the total questionnaire from a coefficient of .84 determined from its halves.

Table 2.—Schematic of Repeated Measures Analysis Design

	Teacher	Discriminating Items	Non-discriminating Items
Lower Class Schools	1	X_{111}	X_{121}
	2	X_{112}	X_{122}
S_1	3	X_{113}	X_{123}
	4	X_{114}	X_{124}
	5	X_{115}	X_{125}
	.	.	.
	.	.	.
	.	.	.
	21	X_{1121}	X_{1221}
Middle Class Schools	1	X_{211}	X_{221}
	2	X_{212}	X_{222}
S_2	3	X_{213}	X_{223}
	4	X_{214}	X_{224}
	5	X_{215}	X_{225}
	.	.	.
	.	.	.
	.	.	.
	21	X_{2121}	X_{2221}

Administration of the Questionnaire

The Basic Skills Questionnaire was given to all the forty-four fourth grade teachers in the lower-class and middle-class schools. Since the Iowa Tests of Basic Skills were given to the fifth grade students at the beginning of the 1969 academic year according to the regular County program, the Basic Skills Questionnaire was given to the fourth grade teachers at the end of the same academic year. Although the students were not necessarily taught by the teachers questioned, it was assumed that the responses to the questionnaire would reflect the general curriculum trend in teaching the basic skills areas given to students in their schools. To get accurate responses from the teachers, their names were not requested. Also, it was stressed that the research would not identify any individual school or teacher. The investigators gave the questionnaire directly to the teachers in twelve schools while the principals, by their choice, gave out the questionnaire in two lower-class schools. Two teachers in one of the latter schools failed to return the questionnaire given out by the principal. Individual school participation was voluntary even though the study had the approval of the County Board of Education Research Committee. With two exceptions, teachers cooperated fully with the study.

Statistical Treatment of the Data

A 2-factor repeated measures analysis design (repetition on one factor) was used as the model in this study (3). Schematic representation of the design is shown in Table 2.

Each of the forty items in the questionnaire was rated by the teachers in the lower-class and middle-class schools. The lower-class schools are designated by S_1 and the middle-class schools by S_2. The notation X_{ijk} denotes the composite score on twenty discriminating and twenty non-discriminating items, where i denotes the school, j denotes the item-category, and k denotes the teacher within the school. Computer time was provided at the Computer Science Center, University of

Maryland, College Park. The following statistical hypotheses were tested at the .05 level of significance: $H_o^1 : \alpha_i = 0$ for all i where α_i is the effect of being in the i^{th} type of school and $H_o^2 : \delta_{ij} = 0$ for all ij where δ_{ij} is the interaction of types of schools and categories of items.

Findings

On the average, the teachers from the middle-class schools estimated the curricular emphasis of their programs to be 3.06 on a 4-point scale. The lower-class teachers estimated their emphasis of the same items as 2.73. A summary of the repeated measures analysis is shown in Table 3. The F ratio observed for the lower-class versus the middle-class schools is 13.59. The critical value of F at 1 and 40 df at the .05 level of significance is 4.08. Hence, the null hypothesis $\alpha_i = 0$ was rejected.

The F ratio observed for the item-by-schools interaction effect was 19.57. Since the critical value of F at 1, 40 df and .05 level of significance is 4.08, the second null hypothesis, $\gamma_{ij} = 0$ was rejected. The fact that there was a statistically significant item-by-schools interaction indicates that the slopes of the curriculum emphasis curves for the lower-class and middle-class schools are different. The slopes of the curriculum emphasis based on the data in Table 4 are shown in Figure 1. This figure shows which category of items were emphasized more by the lower-class and by the middle-class schools.

While there was only a mean difference of 2.67 between the lower-class and the middle-class groups in non-discriminating items, there was a difference of 11.20 in the discriminating items. Figure 1 shows that the non-discriminating items were emphasized in the lower-class schools almost as much as in the middle-class schools. But discriminating items were emphasized very little in the lower-class schools in contrast to the emphasis given these items in the middle-class schools.

Table 3.—Summary of Repeated Measures Analysis on the Basic Skills Questionnaire#

	df	SS	MS	F
Among Teachers				
Lower Class vs. Middle Class Schools	1	1008.11	1008.11	13.59*
Teachers within Schools (Error between)	40	2966.72	74.17	
Within Teachers				
Discriminating vs. Nondiscriminating Items	1	3458.59	3458.59	177.48*
Interaction: (Items x Schools)	1	381.44	381.44	19.57*
Items x Teachers within Schools (Error within)	40	779.48	19.49	
Total	83	8594.32		

*Null hypotheses rejected at the .05 level of significance.

#The assumption of homogeneous variances was supported by the application of F_{max} test.

Table 4.—Mean Composite Estimate of Teachers from Lower Class and Middle Class Schools on Discriminating and Non-Discriminating Items from Basic Skills Questionnaire

School	Discriminating Items (N=20)	Non-discriminating Items (N=20)
Lower Class	48.80	65.90
Middle Class	60.00	68.57

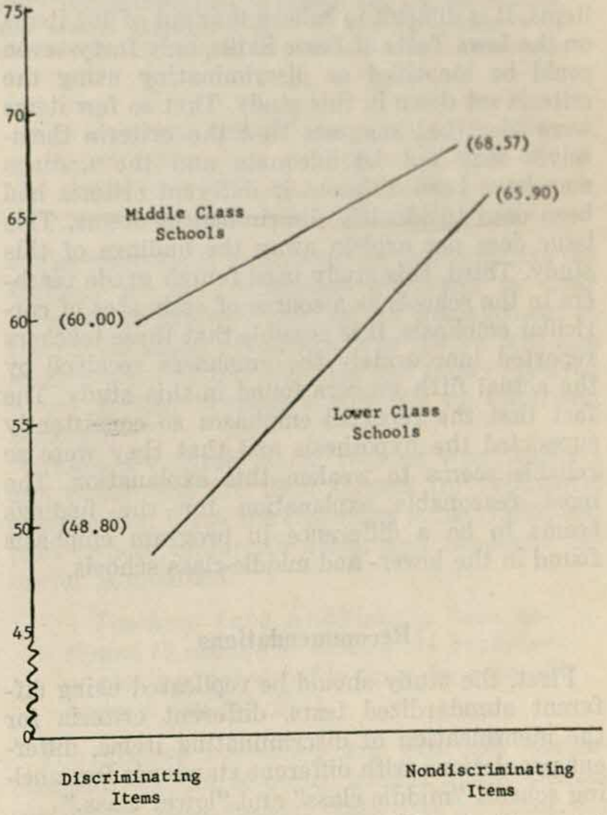


Figure 1. Profile of Interaction Between Items and Lower Class and Middle Class Schools

Conclusions and Discussion

The conclusions, based on the findings of this investigation, are as follows: (1) The research hypothesis, which stated that within the same school system, teachers in lower-class schools report different curricular emphasis in terms of the time spent on the items in the skill areas than do teachers in middle-class schools, was supported by the data. (2) The research hypothesis, which stated that the differences in emphasis reported by teachers in lower-class schools and teachers in middle-class schools are not similarly reflected for the items that discriminate and those items that do not discriminate between the two student groups, was supported by the data.

In examining the findings, at least three rival explanations compete with the research hypotheses in the context of this research. First, it is possible that the differences found between the lower-class and middle-class schools in terms of curricular emphasis are a function of the self-report nature of the Basic Skills Questionnaire. Perhaps teachers in the lower-class schools underrated their emphasis on the skills and teachers in the middle-class schools overrated their emphasis on the basic skills. Although a possibility, this systematic result seems very unlikely. Second, the interactions found between type of school and the discrimination of the items might have resulted from a faulty identification of discriminating items. It is difficult to believe that out of 461 items on the Iowa Tests of Basic Skills, only forty-seven could be identified as discriminating using the criteria set down in this study. That so few items were identified suggests that the criteria themselves may not be adequate and the findings may have been different if different criteria had been used to identify discriminating items. This issue does not explain away the findings of this study. Third, this study used fourth grade teachers in the schools as a source of estimates of curricular emphasis. It is possible that these teachers reported inaccurately the emphases received by the actual fifth graders found in this study. The fact that the reported emphases so consistently supported the hypothesis and that they were so reliable seems to weaken this explanation. The most reasonable explanation for the findings seems to be a difference in program emphasis found in the lower- and middle-class schools.

Recommendations

First, the study should be replicated using different standardized tests, different criteria for the identification of discriminating items, different populations with different standards for labeling schools "middle class" and "lower class."

Second, while the reasons for the observed differences were not identified by this study, steps must be taken by school personnel to eliminate such discrepancies in school program. Use of computer-assisted instructional packages, television programs, and other similar automated programs paced by student abilities and not the whims of teachers may help to decrease the differences in emphases found between the lower- and middle-class schools.

Third, every school system should have item analyses performed on standardized achievement tests so that teachers and parents will know not only that a given student is at or below grade level on a given subtest, but also what particular skills need to be worked on to increase the probability of his succeeding in the American school system.

Fourth, one policy that seems to lie at the root of the problem is that of organizing schools at the elementary level on the basis of neighborhoods. Such a practice tends to segregate schools on the basis of social class. As a result, the "better" students seem to get the "better" program. One way to put a stop to this practice is to establish community schools to serve children of all social classes. The chances are that a lower-class child in a classroom with a lawyer's son and a doctor's daughter may have a more balanced program than if he were isolated in a lower-class school.

All of the foregoing does not assume that all children are equal or that a quality program will make them equal. Obviously there are differences in children's abilities to think abstractly, to paint, to sing, to dance, and to create. However, there is no reason that most school children cannot reach a high degree of mastery in the area of the basic skills. Until this is done, the concept of equal education is a hollow and empty phrase.

REFERENCES

1. Sunnyuh Shin Chang, "Differences in Emphasis on Skills in Programs Directed Toward Children of Varying Socioeconomic Levels," unpublished PhD thesis, University of Maryland, College Park, 1970.
2. Lindquist, E. F.; Hieronymus, A. N., *Manual for Administrators, Supervisors, and Counselors: Iowa Tests of Basic Skills*, Houghton Mifflin, Boston, Massachusetts, 1964
3. Winer, B. J., *Statistical Principles in Experimental Design*, McGraw-Hill, New York, 1962.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 6, February 1971)

Operant Conditioning in the Classroom Setting: A Review of the Research

K. I. ALTMAN and T. E. LINTON
University of Wisconsin–Milwaukee

ABSTRACT

The purposes of this review were (a) to summarize the research following an applied behavioral analysis which has relevance for public school classrooms and (b) to discuss the problems and important considerations involved in an applied behavioral analysis of classroom behavior. Studies in this area are grouped into the following categories according to the nature of the reinforcement employed and assessed: (a) teacher attention, (b) peer attention, (c) token reinforcement, and (d) vicarious reinforcement. The problems of obtaining teacher cooperation, valid and reliable measurements, and confirmation of stimulus control through design of the experiment were discussed. Finally, the implications of these studies for classroom behavior control, as well as for future research, were explored.

WITHIN the last 2 decades several psychologists have moved from an experimental analysis of behavior toward a concern for the applied implications of their laboratory work. This move might be viewed as one from a laboratory study of operant conditioning to behavioral engineering in various community institutions. Both the operant conditioner and the behavioral engineer are cognizant of the theory and applied principles of behavior modification. The operant conditioner, however, is concerned with the advancement of the basic laboratory science and is interested in variables that relate to behavior, while the behavioral engineer has left the laboratory in order to make systematic applications of the principles of operant conditioning to problem behaviors as determined by society (19). This change in direction has been accompanied by a new research strategy, i.e., the applied analysis of behavior and a new body of knowledge.

Some of the current dimensions of an applied behavior analysis have been described by Baer, Wolf, and Risley (4). Many of these dimensions are common to an experimental analysis of behavior as well, e.g., the quantitative characteristics of the target behavior and an experimental manipulation to determine the variables responsible for behavioral change. The major difference between the strategies of experimental and applied behavior analyses lies in their focus.

> *Non-applied research is likely to look at any behavior, and at any variable which may conceivably relate to it. Applied research is constrained to look at variables which can be effective in improving the behavior under study (4:91).*

The purpose of this review is to summarize the findings of the behavioral engineers in their attempts to improve contingency management in the public school classroom. The review will also present some of the problems and issues involved in an applied behavioral analysis of classroom behavior.

There are three major reasons why teachers and classrooms are ideal targets for behavioral engineering. First, the classroom is the traditional place where children's social and academic behavior is modified.

> *Teachers have traditionally been assigned to change or modify the behavior of children entrusted to them for several hours a day (37:288).*

The desired objectives are generally accomplished by the use of praise, grades, and criticisms contingent upon student classroom behavior. By definition, the teacher is a contingency manager, i.e., the teacher knowingly or not reinforces behavior. Given time and experience, most teachers determine which techniques are most effective for them in managing classroom behavior (15). Such trial and error learning by teachers is not only unnecessary but inefficient and often unsuccessful. Writers such as Bergan and Caldwell (7) and

Homme (18) have suggested that the behavioral engineers have already developed a technology which could provide a more effective and empirically based approach to the school-learning process. The latter is probably an overstatement of the present status of behavioral engineering theory. But whether or not it is an overstatement, the theory and technology of behavioral engineering provide a means of developing a more effective and scientific learning model for the public schools.

The schools are ideally suited for behavioral engineering in that the necessary environmental controls are feasible in this setting. That teachers could become contingency managers and would be able to reinforce the desired, and recognize and reinforce approximations to that behavior, however, is not a sufficient condition (19). The contingency manager must be able to set, specify, and to some extent control the environmental variables so that a response by the learner can be reinforced or not as the occasion demands. Public school classrooms are amenable to such control since the students are captive audiences in what is perhaps an already overly controlled environment.

Early Identification and Prevention of Learning Disorders

Another reason for the application of behavioral analysis to the classroom is that these techniques offer much promise for the prevention of those behavior problems which have traditionally led to special class placement and eventual rejection by the school system. Studies by Cohen (11) and Martin and others (27) have shown that some students are not susceptible to the behavioral controls which are usually employed in the public school classroom. However these students can be shaped so that they are able to respond to the usual classroom reinforcement techniques.

The extent of shaping necessary for the disadvantaged and the handicapped child would be greatly reduced if early identification procedures were utilized. Studies have indicated that early identification is possible (38). An alternative approach to prevention lies in utilizing more effectively, or extending the behavioral controls usually exercised in the classroom, so as to deal with a wider range of behaviors.

Emphasis on Classroom Behavioral Control

This review is primarily concerned with the management of those student behaviors usually considered as prerequisites for classroom learning. The concern here is not with the nature and structure of the material to be learned. For this reason, programmed instruction has not been included. The research in that area has recently been reviewed by Leib and others (23). Those studies concerned mainly with special class, institutionalized, or preschool populations will be presented only when they have relevance for the public school classroom. By relevance is meant the presentation of new techniques or the underscoring of old techniques which upon empirical assessment have shown promise for effective application in the usual classroom setting.

For the sake of orderly discussion, it appeared worthwhile to group the studies under the headings *Research On Classroom Contingency Management* and *Teacher Attention*, based upon the nature of the reinforcement contingency employed and assessed; teacher attention; peer attention; token reinforcement; and vicarious reinforcement.

Teacher attention in the form of praise, smiles, and reprimands, when made contingent upon the behaviors of students in a classroom, may serve to increase the rate those behaviors are emitted regardless of the nature, positive or negative, of that attention. Studies of nursery school children summarized by Baer and Wolf (3, 4) have lent support to this assumption and have subsequently led to its investigation in public school classrooms.

Holmes (16) reported the case study of a boy who chronically disrupted the classroom, initiated fights, and was under-achieving academically. An evaluation of the situation indicated that the teacher was reprimanding the child in the classroom and that this served to maintain or accelerate the deviant behavior. A reversal of the strategy by the teacher was initiated. The teacher punished the child by isolating him from the class and attended to appropriate behavior while the child was in the class. This procedure resulted in an apparent elimination of the student's problem behavior in the classroom. As Holmes has indicated, this study has little value other than providing an illustration of the principles of behavior modification in action in a school setting. Since systemic recording of behavior was not carried out by Holmes, we cannot be assured that change did occur.

A preliminary study by Becker and others (6) demonstrated a reliable technique for recording teacher-student behavior in the classroom. Ten children whose behaviors could reliably be rated by trained observers, and who also displayed frequent problem classroom behaviors were selected for study. The children were drawn from five different classrooms each with a different teacher. Following the 5-week baseline period, teachers were instructed to provide rules of good classroom conduct, dispense reinforcement, e.g., smiles and praise, only for behaviors compatible with learning and teaching, and to generally ignore behaviors incompatible with these goals. When teacher approval was made contingent upon appropriate behavior, the rate of such behavior increased significantly. The average deviant behavior for the ten children during the baseline period was 62.13 percent while the deviant behavior during the experimental phase was 29.19 percent. The follow-up study (26) attempted to institute controls which were lacking from the earlier study by Becker and others (6). By instituting rules, praise, and ignoring conditions separately, and employing a reversal to baseline, it was possible to determine the relative effectiveness of each condition by comparing them against a common baseline. Attending to the rules and ignoring inappropriate behavior exerted no substantial decremental effects upon the base rate of inappropriate behavior, but when praise for appropriate behavior was added, inappropriate behavior decreased significantly.

Further support for the effectiveness of praise contingent upon appropriate and not inappropriate classroom behavior in decreasing the rate of disruptive classroom behavior has been gathered by Carnine and others (10). The rate of disruptive behavior was assessed for students when the teachers used only disapproval (baseline) and approval for appropriate behavior and ignored inappropriate behavior for one group of students (B), but did not use approval for another group (A). They then withdrew praise for group A, and instituted the prior group A condition for group B. The dispensing of praise in the classroom was found to affect decreases in disruptive behavior only for those students for whom it was response contingent at that time. Approval, consequated with appropriate behavior, and the absence of approval, consequated with inappropriate behavior, appeared to be the necessary conditions for maintaining classroom behavioral control in this case.

Of those studies discussed thus far, only Madsen and others (25) and Carnine and others (10) made any attempt to reverse experimental conditions. Neither of these reversals consisted of switching the positively reinforced behavior. If teacher praise is indeed a positive reinforcer, it should increase the rate of whatever behavior it consequates. Hall and others (14) provided a test of that notion. From three different classrooms, a total of five students were selected by their teachers as manifesting problem behaviors. After base rates for student study and non-study behaviors were established and teacher attention to each was tabulated by reliable observers, teachers were instructed to pay attention only to study behaviors. This procedure resulted in an increase in study behavior for all three classes.

The next experimental condition called for a reversal in the form of attending only to non-study behaviors. The results of this contingency reversal revealed consequent decrements in the study behaviors and, of course, increases in non-study behaviors of the students. These data then confirm the hypothesis that attention functioned as a positive reinforcer for whatever behavior it followed.

The studies mentioned up to this point have dealt with problem behaviors in public school classrooms with typical pupil-teacher ratios. None of these studies has attempted to assess the effectiveness of teacher attention on all students in a single classroom. Instead, the target population of children with problem behaviors was selected for study. In one of a series of experiments reported by Hall and others (15), study and non-study behavior of an entire first-grade class and consequent attention to each of these conditions by their teacher was recorded. Following the establishment of a base rate, the teacher was instructed to increase positive comments following study behavior, and a concomitant increase in study behavior occurred. Study behavior sharply decreased in a brief reversal when almost no reinforcement for it was provided. The return of reinforcement brought the study behavior back up to a rate slightly above that for the first reinforcement condition. Up to this point, negative comments following non-study behavior were only assessed but not manipulated. Such comments, when made contingent upon non-study behavior remained nearly constant throughout all prior conditions. Following the second reinforcing condition, the teacher was instructed to discontinue all comments following non-study behavior, but to continue her positive comments contingent upon study behavior. As a result, no noticeable change in study behavior or consequently non-study behavior occurred relative to the rates present during the prior reinforcement period. These results shed doubt upon the interpretation that it is merely the amount of attention which is the most important variable influencing the amount of study behavior. If amount of attention alone was the determining factor, a deceleration in study behavior would have been expected when the comments following non-study behavior were discontinued.

A second study which focused upon the entire class rather than target students is that of Madsen and others (26). This study systematically investigated the contention made by Holmes (16) that some behavioral consequences that were regarded by the teachers as punishing or negative may, in fact, have been positive reinforcers for the behaviors they were intended to punish. Prior to this experiment, these teachers had unsuccessfully tried several methods to lessen classroom talking and out of seat behavior for their forty-eight "lower primary" students. The teachers recorded the number of times the students were out of their seats as well as the frequency of "sit down" commands and praise they gave for appropriate behavior. After these base rates were established, a condition was initiated where the teachers were asked to triple the frequency of their "sit down" commands. This experimental condition was conducted twice. Following the second condition of frequent sit-down commands, teachers were instructed to give praise for in-seat and ignore out-of-seat behaviors. Each time the frequency of sit-down commands was increased, the out-of-seat behavior increased relative to the base rates. But when praise was made contingent upon appropriate behavior and inappropriate behavior was ignored, the out-of-seat behavior decreased to a level below both of the prior base rates. Technically, these results do not indicate that punishment accelerated the frequency of the behavior it was intended to decelerate. Instead, these data point out that what is considered a punisher by the reinforcing agent may actually be a positive reinforcer for the student. Theoretically, if sit-down commands originally became aversive by being paired with a primary punishment, they would lose their aversiveness when such punishment no longer was paired. In order to keep the sit-down command enforced, periodic pairing of the commands with aversive stimulation may have to be accomplished.

Each study mentioned has attempted to remove or decelerate behaviors of particular students or entire classes designated as a problem by the teacher involved. Such deceleration was said to have been achieved through systematically varying teacher behavior. It could be argued that such techniques are effective only for restricted populations, i.e., those where "problem" behavior exists. If true, such a criticism would only curtail the generality of the principles involved and not render them invalid or unreliable. But the validity of this criticism is quite doubtful in light of a study by Thomas, Becker, and Armstrong (35). Here, by systematically varying teacher disapproval (e.g., physical restraint, verbal reprimands) and approval (e.g., smiles, praise), an initially "good class" became disruptive. Base rates of teacher approval and disapproval and student disruptive behavior (e.g., noise making and being out of seat) and appropriate behavior (e.g., reading) were established. Then the condition of no approval (i.e., only disapproval) was instituted followed by a base rate. These conditions were then followed by those of no approval, frequent disapproval, no approval, and base rate. Such manipulation suggested that the consequence of withdrawing teacher approval and increasing teacher disapproval results in a concomitant increase in disruptive student behavior. Although the frequency of disruptive behavior was highly variable under most conditions and two independent variables were simultaneously manipulated, the fact that disruptive behavior reached new heights under frequent disapproval is of interest. This finding and those of Hall and others (15), Holmes (16), and Madsen (26) support the tenability of the hypothesis that "negative" consequation in the form of teacher disapproval may function as a positive reinforcer for some disruptive classroom behaviors.

In retrospect, studies of the systematic variation of teacher attention appear to indicate: (a) that teacher approval functions as a positive reinforcer for those student behaviors it is contingent upon; and (b) teacher disapproval, physical or verbal, may also serve as a positive reinforcer or have no effect upon those behaviors, usually disruptive classroom behaviors, with which it is made contingent upon.

Peer Attention

Thomas and others (35) have suggested another interpretation of the increases in disruptive behavior following withdrawal of teacher approval. It may have been that such behaviors were positively reinforced by peer attention. Patterson and others (32) offer some support for the effectiveness of the classroom peer group in controlling the behavior of a classmate. Comparing this study with an earlier one (31), Patterson speculated that "the one variable which seems to account for the difference (a more dramatic effect) among studies is the greater involvement of the peer group. . ." (32:224). Although these studies both effected acceleration in hyperactive children's appropriate behavior in a classroom situation, peer group influence was not systematically studied.

One such study has appeared in the literature, i.e., Evans and Oswalt (12). Two students were selected from fourth grade spelling and arithmetic classes and one student from fifth and sixth grade social-science classes. All of the Ss were considered academic under-achievers by their teachers. For all of these experimental students and their classmates, used as controls, discontinuance of regular classroom activities in the form of story reading or early class dismissal was made contingent upon gains in weekly test performance for the experimental Ss.

Performance gains were assessed relative to base rates and classmate controls. Overall analysis showed that for the six experimental Ss mean gain was significantly greater than that of their classmates. Although this study tends to support the contention that peer influence may have positive reinforcing effects for the behavior of individuals, it is possible that access to the reinforcing events alone was the influential variable. Future studies should contrast the effectiveness of group and individual performance, with individual reinforcement contingent upon individual performance[1]. Such an analysis would clarify whether peer influence is an effective variable in affecting individual performance in the classroom. No conclusions regarding the efficacy of peer influence seem warranted.

Token Reinforcement

Many public school classrooms have traditionally been operated so that the behavioral consequences set up to maintain appropriate behavior are social reinforcement, e.g., in the form of teacher attention, and token reinforcement. Token reinforcement has been defined as ". . . tangible objects or symbols which attain reinforcing power by being exchanged for a variety of other objects such as candy and trinkets which are back-up reinforcers" (30:637). For some, and perhaps most children, doing well in school signified by stars, grades, and points is already a desirable consequence when they enter school. For others they may become incentives by being exchanged for other reinforcers such as access to responsible positions, earlier dismissal from school, and teacher or parental approval. But there may be some children for whom none of the usual reinforcers found in public schools are effective. It has been found that for delinquents (11), retardates, and conduct problem children "generalized social reinforcers such as praise, teacher attention, and positive feedback, often have minimal effects or negative effects upon their classroom behavior" (28:343). For these children, it appears that the traditional social reinforcement is inadequate to establish and/or maintain appropriate behavior. Suggestions for improving the effectiveness of their education is necessary and the research which follows may possess that potential.

Bushell, Wrobel, and Michaelis (9) investigated the effectiveness of token reinforcers in accelerating the study behaviors of preschoolers. Access to special events was indirectly made contingent upon study behavior by providing tokens required to purchase such events directly contingent upon study behaviors. During the second phase tokens and praise were still dispensed for study behavior, but the purchasing power of the tokens was eliminated by permitting all children to engage in special events regardless of the amount of study behavior displayed. The contingent phase was then reinstituted. For ten of the twelve children studied, the results indicated that study behaviors were highest when the special events were contingent and lowest when they were not contingent. It should be noted that only the value of the tokens changed in the non-contingent phase, indicating that the social reinforcement paired with the tokens throughout all phases could not alone maintain or achieve the same rate of study behavior reached by contingent redeemable tokens plus social reinforcement. In the public school, tokens and special events such as museum trips and access to snack times are available reinforcers which usually are not made contingent upon student performance in any systematic manner.

In the Becker and others (6) study discussed earlier, differential social reinforcement from the teacher for appropriate classroom behavior did not seem to effect behavior change in one student until the addition of a daily half-hour of remedial reading. To assess the effect of remediation, Thomas and others (36) observed the frequency of disruptive behavior of a single child in a public school classroom under the following (a) ongoing conditions (87.6%), (b) low teacher approval (69.1%), (c) high approval (51.1%), (d) low approval (64%), (e) high approval and 6 weeks of tutoring (33.4%), and (f) throughout an 8-week follow-up period (11%). The numbers in parentheses show the percent of observation intervals during which the S displayed disruptive behavior. It appears that although teacher approval of non-disruptive behavior decelerated the rate of disruptive behavior, tutoring was instrumental in reducing it even further. It is important to note that the tutoring procedure employed entailed the use of redeemable token reinforcers, marks on a card, paired with praise for academic progress in the tutoring session. Thus, further deceleration in disruptive behavior may have been due to an increase in the efficacy of social reinforcers by virtue of being paired with "back-up" reinforcers. An expansion of the student's academic repertoire thereby increased his potential for emitting appropriate responses and consequently also increased the probability that he will be reinforced socially in class, or both. Although the validity of these interpretations is an empirical question as far as the results of the Thomas and others (35) study are concerned, each provides a different approach to remediation and each may be differentially effective with different students' problems.

Before continuing with the discussion of the efficacy of token reinforcement in controlling classroom behavior, it is important to answer the question regarding the feasibility of this technique in public schools which has been posed by Quay and others (33). Bushell and others (9) have demonstrated the feasibility of token reinforcement in a preschool classroom setting. O'Leary and Becker (30) demonstrated a significant reduction in the "daily mean of deviant behavior" from a baseline range of 66 to 91 percent to a range of 3 to 32 percent during the token program. The class had one teacher and seventeen 9-year-old boys who had been assigned to the class because they exhibited such "undesirable" classroom behaviors as temper tantrums, uncontrolled laughter, and fighting. During the token period each child received points, ranging from 1–10, in a booklet and approval which reflected the extent to which the child had exhibited the appropriate behaviors was listed on the blackboard. Points could be redeemed for a variety of small prizes. The teacher doled out ratings from 3 to 5 times daily at the end of the lessons. The teacher's time involved in giving the ratings never required more than 3 minutes. In fact, the authors noted that the reduction in deviant behavior during the token phase concomitantly increased the amount of time the teacher had to provide more individual attention. This may have contributed in an unknown degree to the lessening of deviant behavior during the token phase. An added feature of this study (30) lies in the finding that the time lapse between the token and back-up reinforcer was gradually extended to 4 days without adversely affecting the level of disruptive behavior. Transfer of behavioral control from tangible tokens to teacher attention adds to the value and feasibility of token systems for use in public school settings. The successful use of a token system in a normal-sized classroom with a more than average number of conduct-problem children answers the criticism of Quay and others (33).

A search of the available literature has resulted in only two studies which have systematically investigated the effects of token reinforcers, as previously defined, on the classroom performance of the entire class in a public school setting. Hall and others (15) were able to increase the mean rate study behavior in a seventh grade classroom with thirty students from 47 to 65 percent by having the teacher increase the frequency of her attention to "appropriate" behavior, and decrease the frequency of her attention to "inappropriate" study behavior. In an attempt to further decelerate inappropriate and accelerate appropriate classroom behaviors, a token reinforcement system was implemented in addition to the previous contingency of teacher attention. The procedure entailed a point system which determined whether the entire class would receive their usual 5-minute break between periods. If any student disrupted the class or was out of seat inappropriately for 5 seconds a chalk mark was placed on the blackboard. Each mark decreased the duration of the break by 10 seconds and twenty-four marks eliminated the break. Study behavior increased to 76 percent under this condition, which might be called token punishment instead of reward. Two days of reversal when break-time was no longer contingent upon behavior resulted in a drop in study behavior of approximately 16 percent. Return of the contingency raised the study rate to 81 percent where it remained until the end of this investigation.

The second systematic study of token reinforcement in a public school classroom was done by Schroer and Johnson (34). This experiment examined "The Effects of Curtailed and Delayed Lunchtime Upon Fifteen Sixth Grade Pupils' Arithmetic Performance." The entire class was subjected to the following experimental phases: baseline, curtailed lunchtime, delayed lunchtime, baseline. In all phases students were presented with ten arithmetic problems to solve in 10 minutes. The problems varied from day to day. During the consequated phases i.e., curtailed and delayed lunchtime, 5 minutes of lunchtime were awarded for each correct problem. The rate of correct problems per minute was .6, 1.2, 1.6, and .8 for baseline 1, curtailed and delayed lunchtime, and baseline 2, respectively. Acceleration of rate correct from baseline 1 to curtailed and delayed lunchtime phases as well as deceleration from these lunchtime contingent phases to baseline 2 were significant. The authors concluded that consequation of arithmetic performance was essential to accelerate and maintain that performance at a relatively high level. It is noteworthy that the increase in rate of incorrect problems per minute was not accompanied by an increase in rate of incorrect problems. Thus, quantity of productivity was increased without sacrificing quality. When academic performance is the target behavior, it is especially important to note whether the gains made were substantial on a pass-fail basis. In this study, only two of the fifteen students remained at a failing level in spite of showing some improvement.

Summary of Token Reinforcement Studies

In spite of the dearth of directly pertinent studies, it appears that public school situations do possess the essential requirements for a token economy system, i.e., potential reinforcing events and point systems to buy access to the necessary reinforcers. Although tokens are employed in the classroom to establish and maintain behavior they are either not dispensed immediately nor are they backed up by potential reinforcing events which are readily available and apparently essential to control the behavior of some students. Events which are reinforcing to students may vary considerably, but the events available should be more than adequate to account for the differences between individual students. Homme (17) and Addison and Homme (1) have demonstrated the value of utilizing reinforcing event menus to establish control of academic and other behaviors. Although some items on these menus are novel they are available in most educational settings. Some adolescent high school dropouts worked through twenty frames of an arithmetic program when consequated with access to working on Russian for 10 minutes (17). The potential of token reinforcement appears to be an as yet largely untapped resource for behavior control in the public school classroom.

Vicarious Reinforcement

Vicarious reinforcement has been defined "as a strengthening of behavior which occurs as a function of observing another being reinforced for that behavior" (10:1). The existence of a vicarious reinforcement effect in a multitude of different laboratory settings has been demonstrated by Bandura (5). Thus far the only attempt to investigate the effect of reinforcing the appropriate behavior of one student in a classroom upon the frequency of such behavior on the other students was made by Carnine and others (10). This study has already been discussed under the section entitled "Teacher Approval." Briefly, that study failed to uncover much deceleration of inappropriate classroom behavior in pupils unless teacher praise, the independent variable, was made directly contingent upon a student's behavior. In fact after contingent praise was withdrawn from Group A and instituted in Group B, increases in disruptive behavior for Group A were more often the case than decreases. Recognizing the inconclusive nature of their findings Carnine and others (10) concluded:

> *The results of the present investigation suggest, but not conclusively, that changes induced by such a procedure (vicarious reinforcement) are either weak or short lived. It is our best guess at this point that praising Tommy serves as a discriminative stimulus for Johnny to behave in a similar way, but that unless some direct reinforcing consequence follows Johnny's improved behavior it will not be maintained.*

An experimental laboratory study by Mithaug and Burgess (29) offers some support for the "best guess" of Carnine and others (10). These authors found that individual feedback was essential to accelerate an individual's response rate in a group situation.

Not enough research has been done to pass any verdict on the efficacy of vicarious reinforcement for classroom behavior control. However, the appeal of such a technique in terms of economy of teacher time certainly warrants future investigation. Even if it does turn out that the effect occurs only when backed-up by occasional contingent attention, the economy of the technique might still be a productive one.

Discussion

Having summarized the findings of attempts by behavioral engineers to improve contingency management in the public schools, we shall now look at some of the problems and important considerations involved in instituting an applied behavioral analysis in the classroom. A discussion of these problems is deemed a more worthwhile pursuit than a premature critique of the research methodology since all of the studies have been recently completed and most are part of long range research programs.

Gaining Teacher Cooperation. In working within a classroom setting several investigators have encountered the problem of gaining teacher cooperation, e.g., Becker and others (6) and Hall and others (14, 15). The following comments from one of the teachers in the Becker and others (6) study concerning cooperation is most indicative of the problem:

> *The plan of the experiment was a bit nebulous, since too much knowledge of what was to be done would affect the results. To add to all this, these people were* psychologists! *My reinforcement history of working with psychologists need not be discussed here (6:288).*

Some of the early investigations in order to obtain cooperation sacrificed some experimental control, e.g., by not using or using only brief reversals (6, 14). Once the teachers were convinced of the effectiveness of the technique and the necessity for reversal, future studies employed reversals, e.g., Madsen and others (25). An additional means of gaining cooperation has been to provide daily feedback to the teachers in the form of graphs illustrating their differential effectiveness under the different experimental conditions. The value of daily teacher-recorded performance rates for their students has been demonstrated by Johnson and others (21). In general, once the teachers realize that such analyses are beneficial to understanding what techniques are most effective in controlling behavior, they in turn are reinforced by the evidence of improvement in the student's behavior. Lindsley (24) has found that it is possible for teachers to successfully utilize behavioral analysis and management techniques after 10 to 20 hours of instruction. The most dramatic illustration of generating teacher enthusiasm for the techniques comes from Hall and others (15). In that study, one teacher was saved from being dismissed because the techniques helped the teacher to gain control of a noticeably disruptive classroom group.

Assessment of Stimulus Control: Validity, Reliability, and Experimental Design. A valid applied behavioral analysis is dependent upon a reliable technique for measuring behavior. Becker and others (6) have presented in detail the steps involved, and the technique they used in assessing classroom behavior. In general, the categories for scoring which fit the situation, i.e., specific or general, and which define the stimulus and the behavioral events in observable terms, are assembled. Observers familiar with the derived scoring categories and their symbols then attempt to use them in the test situation to determine their feasibility and reliability. Agreement between independent observers is taken as evidence of the technique's reliability. Disagreement may indicate that the code is invalid or that different categories are not mutually exclusive. Disagreement may also indicate the faulty training of observers. Inadequacy in the method of assessing agreement is measured in terms of total frequency of behavior instead of the frequency or occurrence on individual segments of the record. Since time sampling is always used, synchronization between observers as to who as well as what behaviors are being observed at a particular point in recorded time is imperative. Time sampling could be done away with if video tapes are used. Every study mentioned in this review employed a time sampling technique. Yet, the assumption of time sampling, that the behavior observed at certain times in a situation is representative of behavior that occurs at other times in that situation, is usually taken for granted. Only Bushell and others (9) made an attempt to assess the validity of the time sample technique by comparing time samples with continuous behavioral ratings. This technique was found to be highly valid in that situation. However, as Bijou and others (8) have noted, such situational idiosyncrasies as frequency of the target behavior necessitate changes in the size of the time sampling interval necessary for valid measurement. The time sampling interval determines the upper as well as lower limits of behavior frequency. For example, if a 10-second interval is used, the maximum response frequency per minute would be six. "Thus in studies with a high frequency of behavioral episodes, small time intervals are employed to obtain high correspondence between the actual and recorded frequencies of occurrences" (8:180). To help insure validity, as well as reliability of measurement, many of these studies hold constant the time and type of tasks performed when manipulations are assessed. Continuing reliability has been insured through periodic spot checks by an additional observer.

The possibility that the teacher or other contingency manager in such situations may also profit from these measuring techniques has been explored by Johnson (20), Johnson and others (21), and Koenig (22). Briefly, these investigations indicate that it is practical and informative for teachers to assess their students' performances in terms of daily rate plotted graphically for target Ss as well as for the entire class. Johnson (20) found that achievement as measured by standardized tests is not the same as when measured by daily performance rates. Daily performance rate would seem to be a more meaningful measure of student performance since it permits immediate and frequent evaluation of progress and is more sensitive to detect specific directions for alternation and/or remediation of curriculum.

Following the method of an applied and experimental analysis of behavior, the majority of studies discussed employed what has become known as an A B A B design. Specifically, such a design entails assessing the base rate, instituting some contingency, removing the contingency in an attempt to return to baseline, and then reinstituting the contingency. Control over target behaviors is said to have been gained when rates of such behaviors reflect manipulations of the independent variables. Replications of such parallel effects are attempted because as Baer and others (4) stated ". . . replication is the essence of believability" (4:95). Aside from the problem of teacher cooperation, reversals sometimes are difficult to obtain because control of the established behavior, for example, by teacher attention, may be maintained by peer attention (2). An alternative design suggested by Baer and others (4) which has not been taken advantage of by researchers in this area entails simultaneous assessment of several behaviors, while "consequating" only one or two at a time. Both of these research designs are powerful in that they depict causal variables and permit analysis both between and within Ss across several conditions. A within-Ss design which does not obscure individual differences is most important when one of the goals of the study is application of the techniques involved. One design that has not yet been employed could be termed an A A B A B design. Madsen (25) mentioned that the training of teachers in the use of operant techniques in the Becker and others (6) study may have caused some target children to improve during the baseline as a result of premature teacher application of the experimental contingencies. Madsen and others (25) precluded such an effect by training teachers after the baselines had been taken. The suggested A A B A B design could assess the effect of training teachers by taking two baselines, i.e., one before and one after teacher training. Such a design may also be more effective in interpreting the results of reversals to baseline attempted later. It may be argued that once a teacher has added these techniques to her repertoire it would be impossible to revert to former traditional teaching methods, and hence, any baseline taken prior to such learning may not be replicable.

Conclusion

In summary, it appears that the entry of the behavior modifiers into the public school classroom has shown that the techniques and designs developed for behavioral control in the laboratory are applicable in natural settings. The ultimate test of the techniques' efficacy will lie in the extent to which they are understood and employed by those in natural settings. As Quay and others (33) have pointed out, there are essentially three major learning situations for the child: the school, the peer group, and the family. This review has concentrated upon the school.

Holmes (16) has illustrated how behavior modification can be carried out when parents and teachers simultaneously apply the techniques. Baer and Wolf (2) and Ferster (13) have pointed out that the key to generalization of the efficacy of a behavior probably rests in the natural community of reinforcement. Certainly a combination of efforts where school, peer group, and family are included in the program for behavioral change would seem a worthwhile goal. As the review has shown, more research is necessary to determine the efficacy of peer influence and vicarious reinforcement in effecting behavioral change. The effects of each of the reinforcement techniques discussed need to be studied on a larger scale. Usually observations account for about 1 percent of a child's behavior during a school day. Videotape recordings, as Madsen and others (25) have pointed out, would extend the scope of the investigation by including more student observations over longer periods of time. This would also provide access to more valid and reliable data gathering.

This review has attempted to demonstrate both the need for and the value of applied behavioral analysis in the public school classroom. It has been demonstrated that teachers can become more effective contingency managers. The immediate and contingent usage of reinforcing events was shown to be a crucial element for effective classroom control. Problems have arisen in securing teacher cooperation, reliable and valid measurement of behavior, feasible and adequate designs and reinforcers, and solutions have either been found or were suggested.

FOOTNOTE

1. This study has been performed since December, 1968, when the present article was completed, cf. Barrish, H.; Saunders, M., and Wolf, M., "Good Behavior Game: Effects of Individual Contingencies for Group Consequences on Disrupted Behavior in a Classroom," *Journal of the Experimental Analysis of Behavior*, 2:119–124, 1969.

REFERENCES

1. Addison, R. M.; Homme, L. E., "The Reinforcing Event (RE) Menu," *National Society for Programmed Instruction Journal*, 5:8–9, 1966.
2. Baer, D. M.; Wolf, M. M., "The Entry into Natural Communities of Reinforcement," paper presented at the meeting of the American Psychological Association, Washington, 1967.
3. Baer, D. M.; Wolf, M. M., "The Reinforcement Contingency in Pre-school and Remedial Education," in Hess, R. D.; Baer, R. M. (eds.), *Early Education*, Aldine, Chicago, Illinois, 1968, pp. 119–129.
4. Baer, D. M.; Wolf, M. M.; Risley, T. R., "Some Current Dimensions of Applied Behavior Analysis," *Journal of Applied Behavior Analysis*, 1:91–97, 1968.
5. Bandura, A., "Behavioral Modification Through Modeling Procedures," in Krasner, L.; Ullmann, L. P. (eds.), *Research in Behavior Modification*, Holt, Rinehart, and Winston, New York, 1966, pp. 310–340.
6. Becker, W. C.; Madsen, C. H.; Arnold, C. R.; Thomas, D. R., "The Contingent Use of Teacher Attention and Praise in Reducing Classroom Behavior Problems," *Journal of Special Education*, 1:287–307, 1967.
7. Bergan, J. R.; Caldwell, T., "Operant Techniques in School Psychology," *Psychology in the Schools*, 4:137–141, 1967.
8. Bijou, S. W.; Peterson, R. F.; Ault, M. H., "A Method to Integrate Descriptive and Experimental Field Studies at the Level of Data and Empirical Concepts," *Journal of Applied Behavior Analysis*, 1:175–191, 1968.
8. Bushell, D.; Wrobel, P. A.; Michaelis, M. L., "Applying "Group" Contingencies to the Classroom Study Behavior of Preschool Children," *Journal of Applied Behavior Analysis*, 1:55–61, 1968.
10. Carnine, D.; Becker, W. C.; Thomas, D. R.; Poe, M.; Plager, E., "The Effects of Direct and Vicarious Reinforcement on the Behavior of Problem Boys in an Elementary School Classroom," unpublished manuscript, University of Illinois, 1968.
11. Cohen, H. L., "M.O.D.E.L.—Motivationally Oriented Designs for an Ecology of Learning," paper presented at the American Educational Research Association Symposium on reinforcement principles to education, February 1967.
12. Evans, G. W.; Oswalt, G. L., "Acceleration of Academic Progress Through the Manipulation of Peer Influence," *Behavior Research and Therapy*, 6:189–195, 1968.
13. Ferster, C. B., "Arbitrary and Natural Reinforcement," *Psychological Record*, 17:341–347, 1967.
14. Hall, R. V.; Lund, D.; Jackson, D., "Effects of Teacher Attention on Study Behavior," *Journal of Applied Behavior Analysis*, 1:1–12, 1968(a).
15. Hall, R. V.; Panyan, M.; Rabon, D.; Broden, M., "Teacher Applied Contingencies and Appropriate Classroom Behavior," paper presented at the American Psychological Association, San Francisco, September 1968(b).
16. Holmes, D. S., "The Application of Learning Theory to the Treatment of a School Behavior Problem: A Case Study," *Psychology in the Schools*, 3:355–359, 1966.
17. Homme, L., "Contingency Management," Newsletter, Section on Clinical Child Psychology, Division of Clinical Psychology, American Psychological Association, 1966, 5.
18. Homme, L., "A Behavior Technology Exists—Here and Now," a version of this paper was given at Aerospace Education Foundation's Seminar, "Education for the 1970's," Washington, 1967.
19. Homme, L.; Baca, P. C'de; Cottingham, L.; Homme, A., "What Behavioral Engineering Is," *Psychological Record*, 18:425–434, 1968.
20. Johnson, N. J. A., "Daily Arithmetic Performance Compared with Teacher Ratings, IQ, and Achievement Tests," unpublished manuscript, University of Kansas, 1967.
21. Johnson, N. A.; Davis, J.; Lindsley, O. R., "Teacher Recorded Performance Rates of Emotionally Disturbed Children," unpublished manuscript, University of Kansas, 1968.
22. Koenig, C. H., "Precision Teaching with Emotionally Disturbed Pupils," unpublished manuscript, University of Kansas, 1967.
23. Leib, J. W.; Cusack, J.; Hughes, D.; Pilette, S.; Wertner, J.; Kintz, B. L., "Teaching Machines and Programmed Instruction," *Psychological Bulletin*, 67:12–26, 1967.
24. Lindsley, O. R., "Theoretical Basis of Behavior Modi-

fication," paper presented at the school of education, University of Oregon, Eugene, May 1967.

25. Madsen, C. H.; Becker, W. C.; Thomas, D. R., "Rules, Praise, and Ignoring: Elements of Elementary Classroom Control," *Journal of Applied Behavior Analysis*, 1:139–150, 1968(a).

26. Madsen, C. H.; Becker, W. C.; Thomas, D. R.; Koser, L.; Plager, E., "An Analysis of the Reinforcing Function of "Sit Down" Commands," in Parker, R. K. (ed.), *Readings in Educational Psychology*, Allyn and Bacon, Boston, 1968(b).

27. Martin, M.; Burkholder, R.; Rosenthal, R.; Tharp, R. G.; Thorne, G. L., "Programming Behavior Change into School Milieux of Extreme Adolescent Deviates," *Behavior Research and Therapy*, 6:371–383, 1968.

28. Meichenbaum, D. H.; Bowers, K. S.; Ross, R. R., "Modification of Classroom Behavior of Institutionalized Female Adolescent Offenders," *Behavior Research and Therapy*, 6:343–353, 1968.

29. Mithaug, D. E.; Burgess, R. L., "The Effects of Differential Reinforcement Contingencies in the Development of Social Cooperation," *Journal of Experimental Psychology*, 6:402–426, 1968.

30. O'Leary, K. D.; Becker, W. C., "Behavior Modification of an Adjustment Class: A Token Reinforcement Program," *Exceptional Children*, 33:637–642, 1967.

31. Patterson, G. R., "An Application of Conditioning Techniques to the Control of a Hyperactive Child," in Ullmann, L. P.; Drasner, L. (eds.), *Case Studies in Behavior Modification*, Holt, Rinehart, and Winston, New York, 1965, pp. 370–375.

32. Patterson, G. R.; Jones, R.; Whittier, J.; Wright, M., "A Behavior Modification Technique for the Hyperactive Child," *Behavior Research and Therapy*, 2:217–226, 1965.

33. Quay, H. C.; Werry, J. S.; McQueen, M.; Sprague, R. L., "Remediation of the Conduct Problem Child in the Special Class Setting," *Exceptional Children*, 32:509–515, 1966.

34. Schroer, P.; Johnson, N. J. A., "The Effects of Curtailed and Delayed Lunchtime upon Fifteen Sixth Grade Pupils' Arithmetic Performance," unpublished manuscript, University of Kansas, 1968.

35. Thomas, D. R.; Becker, W. C.; Armstrong, M., "Production and Elimination of Disruptive Classroom Behavior by Systematically Varying Teacher's Behavior," *Journal of Applied Behavior Analysis*, 1:35–45, 1968 (a).

36. Thomas, D. A.; Nielsen, L. J.; Kuypers, D. S.; Becker, W. C., "Social Reinforcement and Remedial Instruction in the Elimination of a Classroom Behavior Problem," *Journal of Special Education*, 2:291–302, 1968(b).

37. Whelan, R. F.; Haring, N. C., "Modification and Maintenance of Behavior Through Systematic Application of Consequences," *Exceptional Children*, 32:281–289, 1966.

38. Zax, M.; Cowen, E.; Rappaport, J.; Beach, D. R.; Laird, J. D., "Follow-up of Children Identified Early as Emotionally Disturbed," *Journal of Consulting and Clinical Psychology*, 32:369–374, 1968.

BOOK REVIEWS—Continued from Page 271

Although only three chapters emerge as outstanding in a chapter-by-chapter evaluation, overall the book has many contributions to make to classroom teachers. Chief among these contributions is the author's discussion of how basal readers might be used as a launching pad to individualized reading experiences. Stauffer's position regarding the merit of using basal readers is not always clear because of his explicit denunciation on the one hand and his apparent advocacy on the other. However, he eventually shows the reader precisely how basal readers might be used to good advantage in a reading program.

Another appealing characteristic of the book is the generous use of quotations from authorities in related fields as well as in the field of reading. Even though certain of the quotations used seem to lack *"goodness of fit"* in the places they are used, they do provide the reader with some excellent commentary relative to the general topic being discussed. In short, each chapter includes an excellent bibliography.

The clear-cut distinction Stauffer shows between group reading activities and individualized reading activities is commendable. This kind of explication is needed in texts for undergraduate methods courses.

Certain to be appealing and helpful to the intended readers are the illustrations of good teaching practice drawn directly from classrooms. The transcriptions of children's and teachers' dialogue that are presented are well chosen and highly informative of the teaching practices being recommended.

Perhaps the major contribution of the book is the author's unflinching position that reading must be taught as a thinking process from the first day of school. It is to Stauffer's credit that he does not discuss approaches that are not allied to his position. He obviously does not recognize them as good teaching, so the reader must go elsewhere to read of them.

The book has a number of organizational and stylistic weaknesses. Chief among these is the interrupted development of ideas that results from the author's intermingling of philosophical and practical commentary. While obviously intended to give the various discussions depth, this twofold approach often becomes more burdensome than enlightening. At times the reader wishes that the author would be either practical or philosophical, but not both on the same page.

Repetition of key ideas is presumably the author's method of stressing his biases. Unfortunately, the frequent repetition of key ideas may be annoying rather than reinforcing for many readers.

It is difficult to defend the treatment of concept development Stauffer offers the reader in Chapter 10. Perhaps he, himself, explains the weakness of the chapter best when he says, *"The most difficult and demanding aspect of writing this chapter was to condense the abundant source material into a few pages."* The attention given to concept development throughout the other chapters would have been sufficient for the purposes of this text.

Finally, some readers may object to the author's effusive rhetoric. He writes, *"Why, then, these elaborate readiness recommendations? Why? Why? Why?"* And he describes the spirit of a Directed Reading-Thinking Activity by saying, *"There is no cringing before the tyranny of a right teacher answer. On the contrary, devotion to seeking the truth fosters the spirit and confidence of a steadfast scholar."*

The reader who is willing to overlook these weaknesses will find a sound philosophical approach to the teaching of reading as well as many practical suggestions for classroom implementation of that approach.

Richard J. Smith
The University of Wisconsin—Madison

FIELD NEWS

Willard E. North
Central Missouri State College

A STATEWIDE MODEL program for the training of skilled leaders in Wisconsin's grade schools is being developed by a team of educators at The University of Wisconsin and four other universities. The purpose of the training program is to help experienced teachers learn how to become more effective in group leadership, curriculum development, and analysis and improvement of teaching. The first participants are expected to enroll in the program next summer.

The University of Wisconsin—Milwaukee, Marquette University, the State Universities at Eau Claire and La Crosse, and the Wisconsin State Department of Public Instruction are also participating in this program on an equal partnership basis with the University. Professor B. Robert Tabachnick, Chairman of the Department of Curriculum and Instruction in The University of Wisconsin School of Education, is directing the project. Other Madison campus faculty involved in the program are Donald N. Lange, Theodore C. Czajkowski, and Ronald Cohen, Assistant Professors of Curriculum Instruction.

It is believed that the proposed program is unique in teacher education in at least two ways: it regards the problem of preparing effective lead teachers as a statewide problem, and it involves school personnel in the actual designing of the program. Two phases make up the model program. In the first phase, teachers get practice experience on the campuses of participating universities. The on-campus phase is followed by an on-the-job classroom experience while teachers stay in touch with university consultants. Each of the five universities in the project will tailor its own training program to the special needs of schools in different areas of the state. Campus programs at Madison and Marquette University, Milwaukee, will have an urban orientation, while programs at state universities in Eau Claire and La Crosse will stress the needs of teachers in rural area schools. The training program for principals of urban elementary schools will also be developed at The University of Wisconsin at Madison.

* * * * * *

GRANTS TO SUPPORT STUDIES that will increase the understanding of the brain and its relationship to behavior have been announced by the Alfred P. Sloan Foundation. The Foundation will contribute to the support of research and training activities in neuro-science at Cold Spring Harbor Laboratory, Massachusetts Institute of Technology, Rutgers University, and the Salk Institute.

Cold Spring Harbor Laboratory, on Long Island, New York, will establish a summer training program in neuro-science for scientists, engineers, and advanced students from diverse disciplines who wish to become knowledgeable about nervous-system functioning and the techniques for studying it. The Laboratory is under the direction of Dr. James D. Watson.

At Massachusetts Institute of Technology, Dr. Hans-Lukas Teuber has organized a multidisciplinary research group. This group will be primarily concerned with the study of neural mechanisms involved in perception and coordinated movement, internal regulation of drives, and learning and recall. Included in MIT's plans is an undergraduate curriculum in neuro-science.

Rutgers University Institute of Animal Behavior, in Newark, New Jersey, has been a pioneer in exploring relationships between simple behavior patterns and physical and chemical events in the brain. The Institute, headed by Dr. Daniel S. Lehrman, has developed a number of techniques and instrumentation for these studies. The Sloan Grant will assist in expanding this research.

The Salk Institute at La Jolla, California, is initiating a new program in neuro-biology with the aid of the Foundation grant. The Institute has done extensive work in molecular biology and cell biology, and is now extending its interests into research in the functioning of nerve cells and nervous systems of simple organisms.

The Foundation's program in support of neuroscience, directed by Dr. Kenneth A. Klivington, contemplates additional grants over the next several years.

* * * * * *

THE DEPARTMENT of Education of the Banaras Hindu University, Varanasi, India, has recently initiated a long-term research project concerned with the measurement and change of professional attitudes of student teachers. The project is funded by the National Council of Educational Research and Training, New Delhi. Dr. S. P. Ahluwalia, Reader in Education, BHC, is serving as the principal investigator. The anticipated completion date of the project is December 1972.

The objectives of the project are twofold: (1) to develop a multidimensional attitude inventory measuring six areas viz, attitudes toward (a) teaching as a profession; (b) classroom teaching; (c) child-centered practices; (d) educational process; (e) pupils; and (f) teachers; (2) to investigate the nature and extent of change in the professional attitudes of student teachers in selected teacher training institutions of Uttar Pradesh, India.

The first phase of the study places emphasis on instrument development. The second phase will be concerned with examining the possibilities of wide scale use and application of the newly developed instrument for measuring attitudinal change of student teachers. Consideration is being given to future studies that will be concerned with the identification of variables that accelerate or decelerate such changes in attitudes and suggest ways and means for making the teacher education programs in India more effective and efficient.

Presently the research team is reviewing related literature and examining the available teacher attitude inventories. The team is also developing a pool of items for use in the proposed bilingual teacher attitude inventory. As soon as the preliminary draft of the inventory has been prepared it will be given to at least fifty experts for critical review and comments. Further information about the project may be obtained from the principal investigator who would appreciate hearing from others undertaking similar projects.

* * * * * * *

FOR THE TWENTY-FOURTH consecutive year, the Survey Research Center of the University of Michigan will hold a Summer Institute in Survey Research Techniques. The Institute is designed to meet some of the education and training needs of men and women engaged in business and governmental research and other statistical work and of graduate students and university instructors interested in quantitative research and the social sciences.

The 1971 Institute will be presented in two 4-week sessions, the first from June 29—July 23 and the second from July 26—August 21. The two sessions may be taken independently or successively. For further information, write to Mrs. Helene Hitchcock, Senior Administrative Assistant, Office of the Director, Survey Research Center, the University of Michigan, Post Office Box 1248, Ann Arbor, Michigan 48106.

* * * * * *

A LONGITUDINAL STUDY of the Barriers Affecting the Pursuit of Higher Education by New York State High School Seniors—Phase II has been published by the Bureau of Reasearch in Higher and Professional Education of the New York State Education Department. This report presents the findings of a 3-year study being conducted by the Bureau. The purposes of Phase II were to: (1) determine if students implement their declared plans, and (2) analyze those factors which cause students to alter their plans.

Money was the most significant factor for the change cited by the students. When confronted with the fiscal realities of pursuing a post-secondary education, students were forced to make a variety of adjustments. Changes observed in their plans, calculated to reduce college costs, were: (1) to enroll in public institutions instead of private institutions; (2) to enter 2-year instead of 4-year colleges; (3) to commute to instead of live on campus; (4) to attend in-state instead of out-of-state schools; and (5) to work part-time instead of not working at all. The study recommends that counselors and parents be frank in their financial discussion with prospective students. With the exception of financial counseling, the respondents seemed to be satisfied with the overall quality of the guidance and counseling they received in high school. However, those not in college felt that counselors preferred to work with college-bound youth. Of these, three out of four reported that they were furthering their education by enrolling in various forms of non-collegiate education, e.g., on-the-job and specialized training programs. They avoided adult education offerings in local high schools and displayed a preference for schooling offered by industry, business, and the military. This finding may indicate their alienation from the traditional educational structure.

Of the respondents attending college, slightly over half are satisfied with college life. Sixty percent found the academic atmosphere more than satisfactory while 17 percent rated the institution between mediocre and totally unsatisfactory. At least one out of three felt that the college administration evidenced little or no concern for students.

In January 1970, a second follow-up study (Phase III) was conducted using the telephone interviews with the respondents. These findings are being prepared for a separate publication. However, copies of Phases I and II are available. Requests for these publications should be directed to the Editor, SHAPE, Bureau of Research in Higher and Professional Education, The State Education Department, Albany, New York 12224.

* * * * * *

THE INTERNATIONAL INSTITUTE for Adult Literacy Methods has been established by the Government of Iran and UNESCO. The functions of the Institute are to provide documentation, research, and training services on methods, media, materials, and techniques for adult literacy. During the first phase of its operation, the Institute will concentrate on documentation. Public and private bodies actively engaged in literacy work, are invited to send full information on the methods and materials they use, attaching specimens not only of primers and reading materials but also slides, charts, tapes, and films, of teachers' guides, and so forth, along with translation of essential parts, and a synopsis of such material as may be needed. Questionnaires have been sent to institutions to enable the Institute's staff to analyze programs and methods and to select specific items which may be of general interest.

The purpose of this analysis is to summarize and disseminate information on such items in a bulletin to be published by the Institute and to be sent to the services and institutions concerned in order to achieve a wider knowledge of those features which have proved to be particularly successful and to be tried on a larger scale. In some countries a correspondent may be appointed to assist the Institute in coordinating the collection of materials and information.

In the second phase of its operation, the Institute will initiate a program of research on literacy problems. It will serve as a clearing house to acquaint researchers with studies undertaken or planned in various parts of the world, as well as initiating research of its own. The initiated research will consist of inviting others to carry out particular projects in which the Institute, in some cases, may actively participate. This type of research will, in most instances, be on a comparative basis.

In the third stage of its activities the Institute will provide to organizers of literacy work, teachers of literacy workers, and producers of literacy material an opportunity for study and training at the Institute. International seminars or special courses will be organized and it will also be possible for individuals to work at the Institute as research fellows.

In January 1970, the Institute began publication of a journal titled *Literacy Discussion* which is published quarterly and is in both English and French. The summer 1970 issue (Volume I, Number 3) is available free of charge to anyone interested in problems of literacy as well as literacy research. Requests for copies should be sent to Dr. J. D. N. Versluys, Director, International Institution for Adult Literacy Methods, P.O. Box 1555, Teheran, Iran.

* * * * * *

EDUCATIONAL RESOURCES INFORMATION Center/Clearing House for Education and Social Studies/Social Science is the newest of the ERIC Clearing Houses. Its general objectives are the same as Central ERIC's as they apply to social studies/social science information. The major function of an ERIC/ChESS is to collect, select, abstract, index, and make available current pertinent documents. ERIC/ChESS's scope includes all levels of social studies/social science education. The social studies teacher, applications of learning theory, research development projects, the needs of special groups, education as a social science, social sciences and the community, and the content of related disciplines are all included.

ERIC/ChESS is jointly sponsored by the University of Colorado and the Social Science Education Consortium, Incorporated. Its Director is Dr. Nicholas Helburn. Dr. Irving Morrissett, Executive Director of the Social Sciences Consortium and Professor of Economics at the University of Colorado, is the Associate Director. Assistant Director is Mrs. Violet Wagner. ERIC/ChESS is located at 970 Aurora, Boulder, Colorado 80302.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

ecutive Editors

LSON B. THIEDE, Professor of Curriculum and Instruction, e University of Wisconsin, Madison, Chairman.

AYNE OTTO, Professor of Curriculum and Instruction, The versity of Wisconsin, Madison.

JOSEPH LINS, Director of Research, Wisconsin Coordinat- g Committee for Higher Education, Madison, Wisconsin.

nsulting Editors

Terms Expire December 31, 1971

WRENCE P. BLUM, Professor of Educational Psychology, iversity of Wisconsin—Milwaukee.

UGLAS R. GROSS, Arizona State University, Tempe.

LLIAM B. McCOLLY, Professor of English, University of uth Carolina, Columbia.

LLARD E. NORTH, Director of Research, Central Missouri ate College, Warrensburg. Field News Editor.

HN R. PALMER, Professor of Curriculum and Instruction d History, The University of Wisconsin, Madison.

HN E. STECKLEIN, Bureau of Institutional Research, Uni- rsity of Minnesota, Minneapolis.

LIA R. VANE, Professor of Psychology, Hofstra University, mpstead, Long Island, New York.

Terms Expire December 31, 1972

RMAN J. BOYAN, Dean, Graduate School of Education, iversity of California, Santa Barbara.

BERT D. BOYD, Associate Professor of Curriculum and struction, University of Wisconsin, Madison.

BERT CLASEN, Instructional Research Lab., University of sconsin, Madison. Book Review Editor.

RROLL L. GONZO, Assistant Professor of Music Education, iversity of Illinois, Urbana.

B W. JERROLDS, Assistant Professor of Education, Univer- ty of Georgia, Athens.

RL KOENKE, Assistant Professor, College of Education, iversity of Illinois, Urbana.

AROLD E. MITZEL, Assistant Dean for Research, College of ucation, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

GER W. BARDWELL, Superintendent of Schools, Ithaca ty School District, New York.

LORES DURKIN, Professor of Elementary Education, Uni- rsity of Illinois, Urbana.

IVER L. HURLEY, Associate Professor of Special Education, shiva University, New York.

ORGE S. JURATA, Coordinator of Reading, Jefferson County hools, Colorado.

MES D. RATHS, Professor of Education and Director, reau of Educational Research and Field Services, University f Maryland, College Park.

. JAY SAMUELS, Associate Professor of Educational Psy- ology, University of Minnesota, Minneapolis.

ORMAN R. STEWART, Professor, Department of Counseling, ersonnel Services and Educational Psychology, Michigan State iversity, East Lansing.

ditors Emeriti

ARRY J. BAKER, 2241 Q Via Puerta, Laguna Hills, Cali- rnia 92653.

WLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 6757.

ARL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

ARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, alifornia 92802.

. M. JORDAN, Professor Emeritus, School of Education, Uni- ersity of North Carolina, Chapel Hill, North Carolina 27514.

LFRED S. LEWERENZ, Director Emeritus of Evaluation and esearch, 4107 Holly Knoll Drive, Los Angeles, California 0027.

ILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, orida 33146.

. R. MEAD, Director Emeritus of Educational Research, Uni- rsity of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 2601.

ESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

Diary No. 288
Date 16-6-7
File No. Lib
Bureau Ednl. Psy. Res

51st Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **MARCH 1971** **NUMBER 7**

ARTICLES

The Implications for Educational Research of Emerging Planning and Management Systems in Higher Education—*G. Benjamin Lawrence; Clifton C. Lovell; Edward H. Lyell* 291
Factors Related to the Effectiveness of Preschool Programs with Disadvantaged Children—*Julia R. Vane; William M. Davis* 297
Timetable in a Technical University—*Abdul Wahab Khan* 300
Perception of Teacher-Educators in India About the Ideal Pupil—*T. N. Raina; M. K. Raina* 303
Forecasting the Academic Achievement of Engineering Freshmen—*Louis Ross* 307
Student Reaction to the Grade Contract—*Hugh Taylor* 311
Self-Perceived Solutions to Adult Problems—*Ralph C. Dobbs* 315
The Effects of Type of Examination Anticipated on Test Preparation and Performance—*A. Ralph Hakstian* 319
Factors Which are Associated with the Principal's Behavior in Encouraging Teacher Participation in School Decision-Making—*Richard A. Gorton* 325
A Study of Personality Adjustment Among Adolescent Children with Working and Nonworking Mothers—*Deane D. Nelson* 328
A Comparative Study of the Personality Structures of Persistent and Prospective Teachers—*Richard R. DeBlassie* 331

DEPARTMENTS

Among Those Present 336
Book Reviews—*Robert E. Clasen* 296
Field News—*Willard E. North* 334
Forthcoming Articles 318
It Says in The JER 290
New Books from the Publishers 310

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistan to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS Supervisor of Editorial Services.

It Says in The JER...

EMERGING PLANNING AND MANAGEMENT systems according to Lawrence, Lovell, and Lyell are being used by higher education administrators to help them in resource allocation and other managerial tasks. Implications for the educational researcher include an increased role in the decision-making process, an expanded and more accessible data base, and a clarification of the priorities and problems of educational research (p. 291).

SUMMER HEADSTART PROGRAMS seem to be of most benefit to children with lower measured IQ's. Vane and Davis' study results suggest that teachers of preschool programs be encouraged to offer more individualized learning opportunities so the needs of children of varying degrees of ability may be enhanced (p. 297).

TECHNICAL UNIVERSITIES need efficient utilization of resources to make the best use of laboratories, workshops, and drawing halls. Khan describes a heuristic method comprising organization of the data, pre-assignment distribution, and assignment heuristics. By the method, it is economically feasible to construct quality timetables for technical institutions (p. 300).

HOW DO IDEAL STUDENT concepts of teacher-educators in India compare with concepts of teachers in the United States? Raina and Raina found that of sixty-two characteristics ranked, a rank-order coefficient of correlation of .76 was obtained between ranks assigned by the two groups of educators (p. 303).

PREDICTION OF STUDENT SUCCESS in a specific academic discipline was studied by Ross. He used the Wherry-Doolittle method to determine the best battery of predictive measures and then checked the regression equation for practical value as a predictive instrument (p. 307).

HOW DO STUDENTS feel about the grade contract? Taylor found that student responses to a 14-item attitude scale showed a generally favorable opinion. Further evaluation showed that the grade contract group, as opposed to the control group, thought grade contracts were more fair than the traditional method of assigning grades (p. 311).

ADULTS IN DECLINING and non-declining communities perceive some problems and solutions in different ways found Dobbs. He concluded, however, that self-perceived problems categorized as health and education were viewed similarly by adults from both types of communities while physical and economic problems were viewed differently (p. 315).

KIND OF EXAMINATION expected does not affect either preparation or actual performance on either objective or essay tests suggests Hakstian. He first studied college students in an education foundations course and did a second experiment with language arts students (p. 319).

THE PRINCIPAL'S personal role orientation and perception of expectations of his faculty and immediate superior were significantly related to his behavior in encouraging teacher participation in school decision-making, indicated Gorton (p. 325).

DO WORKING MOTHERS affect the personality adjustment of their children? Nelson found that personality adjustment of boys was better if the mother worked full-time than if she worked part-time or not at all. There was no consistent pattern for girls, but on the majority of MCI scales girls with non-working mothers had better adjustment scores (p. 328).

SLIGHT PERSONALITY differences exist between teachers with undergraduate teacher training and prospective teachers with undergraduate liberal arts background. DeBlassie studied teachers and teachers-in-training using the Sixteen Personality Factor Questionnaire (p. 331).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Diary No.
Date
File No.
Bureau Ednl. Psy. Research

Editorial/Abstract

The State of the Art:

The Implications for Educational Research of Emerging Planning and Management Systems in Higher Education

◆

In addressing the many problems facing higher education today, administrators have turned to emerging planning and management systems to help them in their resource allocation and other managerial tasks. The development of these planning and management systems—given present trends—will also provide benefit to the educational researcher. The benefits will be increased in proportion to the degree of cooperation and involvement of the educational research community in the development of these emerging systems. The implications for the educational researcher include: an increased role in the decision-making process, an expanded and more accessible data base, and a clarification of the priorities and problems of educational research.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

The Implications for Educational Research of Emerging Planning and Management Systems in Higher Education

G. BENJAMIN LAWRENCE, CLIFTON C. LOVELL, and EDWARD H. LYELL
Western Interstate Commission for Higher Education

IN ATTEMPTING to deal with the myriad of problems common to most colleges and universities today—ranging from the allocation of limited resources, to student unrest and public dissatisfaction with the perceived benefits of higher education—top level decision-makers are turning with vigor to the development of tools and procedures they believe will help them improve the management of their institutions.

These emerging planning and management systems have distinct implications for the educational research community based on the following assumptions and constraints:

> *1. That these systems deal with higher education as opposed to elementary and secondary education. While the authors believe that some of the observations which follow may be generalized to all educational levels, the situations are not analagous. What is discussed here in relation to higher education should not indiscriminately be generalized.*
>
> *2. That the presently emerging planning and management systems will, in fact, continue to be developed.*
>
> *3. That this development continues in accordance with current philosophies, assumptions, and procedures. A major shift in the assumptions on which such systems are currently being developed may markedly change their anticipated future impact on educational research.*

The underlying concepts in the development of these new planning and management systems suggest at least three major implications for educational research:

> *1. Educational research will play an increasing role in the decision-making processes of higher education in the future.*
>
> *2. The data base upon which educational researchers may draw will be expanded and made more accessible.*
>
> *3. Priorities and problems for educational research will be more clearly explicated given the clearer statement of educational objectives and goals (benefits and products) motivated by the implementation of planning and management systems.*

This paper will attempt to address each of these implications in terms of its impact on the educational research community.

The First Implication

An Increased Role for Educational Research in the Decision-making Process. Borrowing from elementary and secondary education terminology, "educational research" as opposed to "administrative research" has, in higher education, tended to be much more academic than operational. Except in a few more enlightened institutions, research on the learning environment and processes has been largely academic, peripheral, and seldom incorporated into the decision-making process. With the exception of class, section, and laboratory size and a cursory review of the subject matter content, the professor is left largely to his own devices with respect to the organization and implementation of the educational process. Educational research may well affect the individual professor but seldom the institution as a whole.

Frequently, the educational research community, and to some extent the entire academic community, seems to be at an opposite pole from the management process. Most institutions have two major divisions within their organizations—academic affairs and business affairs. The academic affairs division is concerned with educational content and process, while the business affairs division is responsible for operations and budget in support of the academic function. They are concerned with means and product on the one hand and efficiency and cost on the other.

While educational researchers, academic planners, and business managers alike ultimately share a common concern over the financial implications of the educational process, the planning, programming, and budgeting procedures utilized in colleges and universities, to date, frequently have not contributed to the integration of the efforts and expertise of these individual functions. In fact, they have tended to compartmentalize these respective activities and emphasize cost and efficiency. This emphasis presented few problems so long as funds were available in sufficient quantity to provide support for the plans as perceived by the academic planners. Institutional researchers accordingly tended to emphasize administrative as opposed to educational research. Administrators needed more information—quantitative in nature—to justify administrative requests for more money to meet increased enrollments and increased demands for new programs. Institutional researchers were forced to use the majority of their limited resources in providing this type of information. Educational research was in little demand when funds were freely available, because better, more efficient ways of improving educational processes were perceived to be of little value.

The process of higher education is viewed from the standpoint of the emerging planning and management systems as shown in Figure 1.

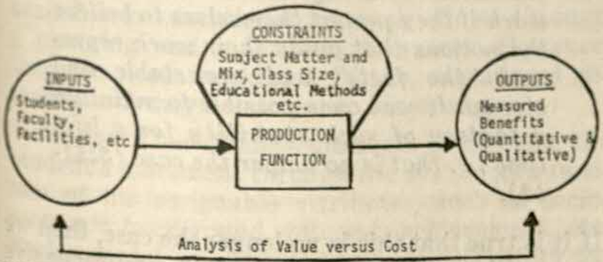

CONSTRAINTS
Subject Matter and Mix, Class Size, Educational Methods etc.
INPUTS
Students, Faculty, Facilities, etc
PRODUCTION FUNCTION
OUTPUTS
Measured Benefits (Quantitative & Qualitative)
Analysis of Value versus Cost

FIGURE 1. Higher Education Process Viewed From Standpoint of the Emerging Planning and Management Systems

Given limited resources and a desire to maximize the benefits of higher education, the necessity for integrated effort on the part of business managers, academic planners, and educational researchers becomes apparent. The increased role of educational research in attempting to improve the production function of higher education then becomes easily identified.

Questions such as the following are generally considered to be topics of investigation by educational researchers: What educational techniques given specific objectives are the most effective in light of available funds? What mix of inputs—facilities, faculty, services, etc.—best enhances learning in what specific discipline? The answers are clearly necessary information for improved management and planning of higher education as described by the emerging planning and management systems approaches.

To explicate more clearly the potential importance of educational research, consider the following hypothetical management problem:

> *Five competing programs—academic disciplines A,B,C,D,E—each submit budget requests to management. Each program specifies the dollars required and the benefits to be derived from the expenditure. While the benefits of all programs are perceived as desirable, the benefits per dollar of expenditure in program C are such that the benefits of supporting program C are not perceived by the decision-makers (president and regents) as being as worthwhile as spending those dollars to purchase more of the benefits in programs A, B, D, and E. (Five passenger cars at $9,000 each might be nice, but one bulldozer and three 5-ton trucks are more effective.)*

Given this situation, an educational researcher may be able to discover how the same product can be produced for less, or he may design a product of different quality and performance specifications which lower the cost/benefit ratio making the benefits of program C more attractive to the decision-makers. (Five passenger cars at $3,000, one additional bulldozer, and one more 5-ton truck may appear more attractive than one additional bulldozer and three 5-ton trucks.)

A Second Implication

An Expanded and More Accessible Data Base. The emerging planning and management systems are predicated on the assumption of an enlarged and readily manipulated data base from which valid and relevant information can be derived for addressing the problems confronted by the decision-makers. The commitment on the part of administrators to the enlarged data base associated with emerging planning and management systems stems from the desire to have better quantitative and qualitative management information, and from the increased perceived value to management decision-making that will result from the findings of educational research made possible by this enlarged data base.

Educational research projects have generally been restricted to *ad hoc* studies of short duration, dealing with a single aspect or problem of education, and without a larger conceptual framework against which to examine results. This is due in part to funding methods and restrictions, and to the transient nature of the student sample population. This situation creates incompatibility of study results, thus inhibiting the development of a uniform body of knowledge. Even worse, the present system tends to encourage short-term studies against which comparisons will not or cannot be made.

In an address to the American Educational Research Association, Goodlad, Dean of the Graduate School of Education at U.C.L.A., noted:

> *Most of the scholarly inquiry in a field of any given time is of a short term variety. Consequently, the intrusion of new conceptual structures into the field is upsetting—particularly to those short-term inquirers who have mortgaged a chunk of their lives in the grant-type of structures now being challenged (2:10).*

While this statement reflects the present state of a large portion of the research community, it must be added that this need not be the case given an adequate information system. Both the conceptual structures and the information necessary to relate them to the short-term inquiry can be a continuing and integral part of such a system. Goodlad in the same address also added:

> *. . . short-term inquirers, although they need not be model-builders, must be aware of the need to check the fit and the parallel with respect to conceptual and operational models lest their research be barren (2:18).*

Ad hoc studies have in their basic nature the problem of selecting and/or recruiting a test group which, in itself, can be bias inducing. This group is then used in the study, and hypotheses concerning education are tested against the group's performance, which is either already recorded in some form such as earnings versus grade point average or is recorded under test conditions.

Of such studies, Dewey employs the term "working hypotheses" and says:

> *. . . they are working ideas; special investigations become barren and one-sided in the degree in which they are conducted without reference to a wider, more-general view. . . . No matter how these are obtained, they are intrinsically philosophical in nature (1).*

The concept of an intelligence quotient (IQ) is a case in point. Without tracing the entire history of this measure and the "statistical" misconceptions about it that have flourished and subsequently fallen by the wayside, we need only recall a few highlights in its history. During the early years of its usage, a number of tests and studies supported the theory that it was an invariant measure of one's learning capacity, whereas environment and other influences affected the degree to which this capacity was utilized. Subsequently, the IQ score was misused in a number of ways, not the least of which were studies concluding that Jews were inherently more intelligent and blacks less intelligent than the population as a whole.

We are now more aware of the pitfalls associated with the IQ measure and the factors that influence its outcome; however, we are not particularly better equipped to prevent a similar failing from occurring again. In fact, the absence of a comprehensive data base structured against an educational framework continues to force us to design *ad hoc* data systems for most studies, providing the same opportunity for "statistical" misconceptions that led us down the primrose IQ path.

The IQ score also provides us an example of the need for continuous data. Not only were the researchers in IQ qualities victims of inconsistent environmental, cultural, and socioeconomic measures across groups, but they also had difficulties in tracing sufficient samples longitudinally. The need for continuous longitudinal data cast against a conceptual framework is critical to the study of long-term effects of education. Schwab uses the term "fluid inquiry" for this kind of study and says:

> *. . . there are some scientists who would deny the very existence of fluid inquiry because they deny the existence of the constructural frames as underpinnings of their work. Many men feel much more emotional stability and readiness for work if they permit themselves to believe the notions that guide their work represent "the facts" and are stable and eternal. It was once possible to maintain a fantasy of such stability for a lifetime . . . that is no longer the case (3:41-42).*

If it is true that this is no longer the case, then we are overdue in the development of comprehensive information systems to assist in the longitudinal study of higher education.

Advantages of the Data Base and Responsibilities of the Researcher

A data base generated for the purpose of supporting comprehensive planning and management systems and the models of the system itself provide direct and tangible advantages for educa-

In particular, the data base would provide information which is:

1. **Continuous and longitudinal** (information over time)
2. **Contiguous and latitudinal** (information across disciplines, student characteristics, etc.)

These data will provide the decision-maker and the educational researcher with information on student and faculty attributes and behavior, as well as organizational and operational characteristics of the institution over time.

The collection and organized analyses of these data will provide the educational researcher with information which he can apply to his research studies, thus permitting him the opportunity to reduce the inferential jump required in comparing data from his specific studies with the actual educational environment. This is made possible through a common base vocabulary and set of definitions used in the data collection process. Even though direct comparisons are not possible, this compatibility of definition will allow identification of the differences between studies and environments.

However, the provisions of such comprehensive and common data will provide an effective base for the educational researcher only if he participates in the definition of the data and the definition of the associated models. Since he, as well as the decision-maker, is an intended consumer of these data, his help is needed if the data are to be relevant to his needs. Therefore, the educational researcher must speak out when he sees the development of these systems varying from his needs, either in completeness, or in type.

Process Models

The following four examples are merely indicative of the kinds of models that can be made available to educational research efforts through a planning and management system. However, they too need the inputs of the educational researcher if they are to serve his needs.

Student Flow Model—attempts to describe the student's movement through the system as a function of his assignable attributes, such as socio-economic background, interest, achievement, etc. Such a model, in addition to its obvious utility for predicting student trends, should provide the researcher with a comprehensive record of student behavior and decision-making as a function of a wide range of contributory variables.

Faculty Activity Model—should provide the researcher with equivalent cognitive information on faculty behavior in an interrelative way, thus allowing analyses on student-faculty interaction.

Induced Course Loading Model—has its roots in managerial analysis of the impact on one in another. The basic concept of modeling induced course loads of this type inherently includes the relative mix of fields and courses associated with a specific degree program. While individual program make-up can be studied with the present systems, the interrelationship of programs and the degree of specialization and non-specialization is retrievable by degree program in a model of this type. (Such information appropriately related to other variables should be very useful to the researcher concerned with program and curriculum effects on student performance. At the least, it provides a frame of reference against which alternatives may be evaluated. A simulation of the effects of various changes could be undertaken with this kind of model.)

Resource Requirements Model—has as its primary motivation managerial decision-making. However, from the managerial standpoint, such models provide insight into the resource impact of alternative programs with no measure of benefit or effectiveness inherent in the design. This model, like the induced course loading model, provides the researcher with a framework in which to measure the effect of any proposed changes resulting from his research findings, in terms of the resource constraints of his institution.

The utilization of planning and management models of the types described should assist the educational research community in the development of an overall conceptual framework against which future studies can be structured. This framework, in conjunction with the emerging data systems, will not only provide significant gains to educational research, but will also provide the vehicle by which management can better choose between alternative results of research leading to a better opportunity for implementation.

A Third Implication

The Clarification of Priorities and Problems of Research. A major thrust of the emerging planning and management systems is to obtain a clear explication of the goals and objectives of higher education in the form of benefits and products (or their proxies) which are measurable. This is a difficult task given the intangible nature of many of the perceived benefits. The state of the art is primitive; nevertheless, serious attempts are being made. A number of studies of benefits derived from higher education already fall in the domain of educational research. As these benefits or products become more clearly explicated, the problems and obstacles hindering their obtainment also become much clearer, providing more precise articulation of the targets of educational research and better understanding of the priori-

Summary

Educational research can provide considerable assistance to the emerging planning and management systems through a better understanding of the less tangible products of higher education. On the other hand, the emerging systems will provide educational research with a vehicle for better identifying the more significant and critical areas requiring the attention of the educational research community as well as a considerably expanded data base upon which to draw for their research efforts. The role of educational research will be more greatly valued and expanded in scope as a result of the impact of emerging planning and management systems on the administration of higher education.

REFERENCES

1. Dewey, John, *The Sources of a Science of Education*, New York, Horace Liviright, 1929.
2. Goodlad, John I., "Thought, Invention, and Research in the Advancement of Education," *The Educational Forum*, 33: (no. 1) 10, 18, November 1968.
3. Schwab, Joseph J., "The Structure of the National Sciences," in Ford, G. W.; Pugno, Lawrence (eds.), *The Structure of Knowledge and the Curriculum*, Rand-McNally, Chicago, 1964, pp. 41–42.

Book Reviews

ROBERT E. CLASEN
Book Review Editor

Psychosocial Principles Applied to Classroom Teaching

White, William F. McGraw-Hill Book Company, New York, 1969, 228 pp.

HERE IS yet another book written by an educational psychologist who claims that his discipline has "enormous" amounts of information that should be presented to prospective teachers. He advocates that teacher preparation programs spend a larger number of academic hours in five subfields of educational psychology of which social psychology is but one.

The author has reviewed the literature in several important areas of social psychology, including the best of the recent and the past. The researches are summarized in an interesting style and in most cases with caution in qualifying the conclusions. The weakness comes when the author tries to list "applications"—the reader will find precepts such as "Demonstrate interest in the subject matter" (p. 27); dicta such as "the teaching role is not merely to impart information but also to satisfy the emotional needs of both the student and teacher" (p. 29); epigrams such as "the teacher's effectiveness is determined by the control he generates in the students who sit in the classroom" (p. 93); and proverbs such as "students with positive and healthy self concepts are more linked than children with negative self-concepts" (p. 160). None of these, nor most of the many others listed in the text seem to be truly applications, defined as "a method of using" by my Webster's.

In sum, the text is a wonderful review of research in the area of social psychology as applied to educational problems. The lists of applications are at best amusing.

Professor James Raths
College of Education
University of Maryland, College Park

Changing Classroom Behavior: A Manual for Precision Teaching

Meachem, M. L. and Wiesen, A. E. International Textbook Company, Scranton, Pa., 1969, 212 pp.

THIS BOOK is the latest in a number that have been published recently that are aimed at helping teachers apply the principles of learning theory in the classroom. According to the authors

> *Precision Teaching . . . blends what we know as educators with what we know as behavioral scientists. We have attempted to provide rather specific guidelines for teachers regarding positive behavior change in the classroom.*

The book is true to its purpose in that it outlines specific principles of learning which can be applied by the classroom teacher and also provides the results of research to give the teacher a more extensive understanding of the basis of the principles of precision teaching.

Part I contains the instructions for precision teaching and is divided into eight chapters. The first chapter sets the stage by challenging the teacher.

> *Revolution is not limited to only politics, social systems of science, but can occur in education as well. . . . Today a revolution is in process which is already beginning to make profound changes in both the theory and practice of teaching. . . . What is developing is a system of education which carefully evaluates its own methods through the framework of science and makes corrections based on data rather than opinion, on facts rather than fiction. In the classroom, this new approach is beginning to manifest itself in what is called Precision Teaching. As it becomes more widely known, Precision Teaching is going to play a major role in the classroom. To you, the teacher or prospective teacher, it means an exciting addition to your teaching skills.*

The six chapters that follow are involved with explanations of how to measure behavior in the classroom, set behavioral goals, strengthen desirable behavior, eliminate undesirable behavior, provide cues for learning, and program for successful learning. A limited number of terms associated with learning theory are introduced and explained and a 3-page glossary is provided at the end of the book.

The eighth chapter covers ethical considerations and

(Continued on page 327.)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Factors Related to the Effectiveness of Preschool Programs with Disadvantaged Children

JULIA R. VANE
Hofstra University, Hempstead, N. Y.

WILLIAM M. DAVIS
Health and Welfare Council, Nassau County, N. Y.

ABSTRACT

This study of a group of disadvantaged children who attended a summer Headstart program indicated that many of the children made gains in tested IQ and that many of the children had higher measured intellectual ability than is usually expected. It is also indicated that children with lower measured IQ's seemed to benefit most from the program, whereas children with higher IQ's showed less gain as measured by a retest at the end of the program. The results suggest that teachers of preschool programs should be encouraged to offer more individualized learning opportunities so that the needs of children of varying degrees of ability may be enhanced.

THERE HAS been considerable concern among educators regarding the fact that many compensatory education programs have failed to show expected results in terms of improvement in children as measured by standardized tests. Bereiter and Englemann have suggested that "radical departures from established practices of early childhood education are needed" (1:19). The present study suggests that greater individualization may be required if all disadvantaged children are to benefit.

Subjects and Procedure

The 371 children who were evaluated attended seven different Headstart Centers, for a 6½ week period during the summer. Collection of the demographic data and initial testing of the children was done during the first 2 days of the program. Follow-up evaluation was possible in only four schools, because of lack of staff. The final testing was done during the next to last day of the program, thus there was a 6-week period between test and retest. Plans for final evaluation were not made known to the staff in the Centers to avoid any tendency to train for retest.

Results and Discussion

A summary of the demographic data and the test results are shown in Table 1. As might be expected, the great majority attending were children of blue collar workers (50%) or children whose parents were on welfare (31%). There were slightly more boys than girls, but an equal number of white and Negro children. The initial testing showed that the mean IQ of the children in the seven Centers was 95.5, which falls within the average range. This level of performance is higher than that usually associated with children of this socioeconomic background. This contradicts the findings of Bereiter and Englemann namely, "With no known exceptions, studies of 3- to 5-year-old children from lower socioeconomic backgrounds have shown them to be retarded or below average in every intellectual ability" (1:4).

Two factors may be responsible for the average IQ level found in the present study. One is that the level of tested intelligence in Nassau County has been shown to be higher than that of the population in general (4). The other is that the test used, the Vane Kindergarten Test (VKT) (2), has been standardized on both white and Negro children of this age group and therefore may more adequately represent their performance level than the usual tests. The test results are felt to be comparable to those that might have been obtained on the Stanford Binet. In three separate studies in which the Stanford Binet and the Vane Kindergarten Test were given to both lower class and middle class children, the mean Binet IQ's and VKT IQ's were not more than 3.9 IQ points apart (2, 3). In the first study of seventy-eight kindergarteners in a middle class district, the Binet mean IQ was 2.7 points higher. In a second study involving 212 kindergarteners in an integrated lower class district the mean Binet IQ was .8 of a point higher. In a third study carried out in a full year Headstart nursery school program with fifty-one children, the VKT was higher by 2.1 IQ points.

Analysis of test scores and demographic data suggests that there was a trend for better scores to be associated with centers where there were more children with parents in white collar occupations and lower scores to be associated with

Table 1.—Demographic Data and Test Results for Seven Preschool Programs

		Sex		Race		Occupation*				
School	No.	M	F	W	N	White Collar	Blue Collar	Welfare	Mean IQ	SD
A	62	45%	55%	73%	27%	8%	63%	18%	95.1	13.5
B	48	48	52	58	42	13	56	29	99.5	11.6
C	75	53	47	5	95	4	31	63	90.9	12.1
D	69	51	49	100	0	35	52	4	102.0	13.3
E	34	65	35	62	38	24	56	21	99.0	11.0
F	33	58	42	24	76	4	56	37	89.9	11.2
G	50	58	42	24	76	12	46	42	98.9	9.0
TOTAL	371	53%	47%	50%	50%	14%	50%	31%	95.5	12.9

*Data not available in every case so figures do not add up to 100%.

schools where a large percentage of the children came from families on welfare. Center G was one in which the trend was reversed. Despite the fact that 42 percent of the children in the Center came from families on welfare, the test scores were higher than other districts in which fewer children came from the lower socioeconomic ranks.

The test retest data, for 116 children, shown in Table 2 indicates that gains were made in all schools in tested mean IQ. The greatest gain was in the center with the highest initial mean IQ and the least gain in the center where the mean IQ was lowest initially. The gains in mean tested IQ are felt to indicate that the children, as the result of their summer Headstart experience, learned to respond more adequately in a testing situation than they had initially. Since a testing situation requires attention, concentration, and an adequate response, improvement in these areas should help children profit from future education.

Although the majority of children showed gains in IQ, some did not. The percentage of those who gained, ranged from 60 percent in one center to 89 percent in another. Analysis of the individual test protocols indicated that there was a trend in most of the programs for the children who had lower IQ's initially to show more gain than those who had higher IQ's initially. This trend was shown in every center, and the data are summarized in Table 3. The difference between the mean IQ of the group that gained and the mean IQ of the group that lost is significant at the .01 level. The results suggest that the teaching in these centers may have been directed toward developing the abilities of the poorer performers while neglecting to enhance the skills of the

Table 2. Results of Pre- and Posttesting at Four Headstart Centers with the Vane Kindergarten Test

					No. of children Showing			
Center	No.	Pretest Mean IQ	Posttest Mean IQ	Mean Gain	Gain	No Gain	Loss	Percent Gain
C	45	90.4	91.7	1.3	28	3	14	62%
F	20	92.5	95.3	2.8	13	0	7	65%
E	24	94.8	98.2	3.4	15	3	6	63%
G	27	99.6	105.6	6.0	23	1	3	85%
TOTAL	116*	93.7	96.9	3.2	79	7	30	68%

* The discrepancy between the enrollment as shown in Table 1 and that shown here reflects the number of children who were not in attendance on both days that the testing took place.

Table 3.—Initial Test Scores of Children who showed Gain, No Gain, and Loss after 6 Weeks Attendance at Summer Headstart

Center	Gain N-79	No Gain N-7	Loss N-30
C	76,76,78,78,78,79,79,79 80,80,80,80,81,82,84,85 86,86,86,88,88,88,90,94 95,95,102,108	80,85,89	87,88,89,90 92,99,102,102 102,105,110,112 121,125
	Mean IQ 85.0		Mean IQ 101.7
F	79,81,83,86,86,88,88,89 89,90,97,100,120		85,91,96,96,96, 100,110
	Mean IQ 90.5		Mean IQ 96.3
E	76,82,84,85,86,87,87,91 92,94,94,104,106,109,111	74,91,101	84,98,101,105, 115,119
	Mean IQ 92.5		Mean IQ 103.7
G	85,86,86,87,88,88,91,91 96,97,97,100,100,100,100, 101,102,107,107,111,112, 112,113	116	99,103,107
	Mean IQ 98.1		Mean IQ 103.0
Total	Mean IQ 91.2 * S.D. 10.15		Mean IQ 101.0 * S.D. 12.15

*Difference significant at the .01 level.

children with more advanced ability. It is possible that the children with higher intelligence who watched the children with more limited ability being praised and given attention, may have found that they received more attention when they did not perform as well. Comments on the individual test protocols and a comparison of the records of a number of the brighter children who showed losses on the retest, indicated that these children appeared more hesitant, less confident, and showed less initiative and imagination on the retest than they had on the original test.

Discussion with the psychologists who were working with the individual programs indicated that they felt that the programs tended to be oriented toward the slower moving children. It appeared that many staff members may have expected disadvantaged children to be unable to function on an average or better level. As a result, they did not require enough from them, nor did they provide materials that were cognitively stimulating for such children.

The fact that the scores of the children at the lower end of the IQ scale tended to move upward and the scores of the children at the upper end of the IQ scale tended to move downward, might suggest that the changes between pretest and posttest scores might be reflecting the phenomenon of regression toward the mean. An analysis of the figures shows that, if the scores are considered on a school by school basis, 82 percent of the children who tested below the mean of the school showed gains, whereas only 52 percent of the children who tested above the mean of the school showed losses. If the group is considered as a whole and a group mean obtained, then 80 percent of the children who were below the mean showed gains, and 44 percent of the children who were above the mean showed losses. In both instances, the difference between the percentage of those below the mean who gained IQ points and the percentage of those above the mean who lost IQ points was significant at the .01 level. This suggests that the difference was not likely to be a result of regression toward the mean for if this were so, approximately the same percentage of scores at either end of the scale might have been expected to move toward the mean.

Although intelligence testing has been criticized because some feel it has led teachers to have lower expectations of some children, the critics do not mention that test results often indicate that many disadvantaged children have more ability than expected. The results of the present study indicate that compensatory preschool programs might be more effective if the programs were planned to meet the needs of children who represent a broader range of intellectual potential than many expect to find in groups of disadvantaged children.

REFERENCES

1. Bereiter, C.; Englemann, S., *Teaching Disadvantaged Children in the Preschool,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1966.
2. Vane, J. R., "The Vane Kindergarten Test," *Journal of Clinical Psychology,* 24:121–154, 1968.
3. Vane, J. R., "Use of the Vane Kindergarten Test," paper presented as part of a Symposium on Rapid Screening Devices, The American Psychological Association Convention, Washington, D. C., 1969.
4. Vane, J. R.; Weitzman, J.; Applebaum, A., "Performance of Negro and White Children and Problem and Nonproblem Children on the Stanford Binet Scale," *Journal of Clinical Psychology,* 22:3, 4, 431–435, 1966.

# # #

OUR SCHOOL systems need more than just safe and trouble-free administrations. They need the challenge of creative thinking. Even though the security of mediocrity may be appealing to the administrator who feels beset on all sides, it is apparent that our schools cannot afford it.—*Benjamin J. Willis.*

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Timetable in a Technical University

ABDUL WAHAB KHAN,[1] **West Pakistan University of Engineering and Technology, Lahore**

ABSTRACT

This paper describes special features of a timetable, in a technical university, such as practical instruction in laboratories, workshops, and drawing halls spreading over 2 to 4 consecutive hours. It emphasizes the need for more efficient utilization of available resources as compared with that in general education. The scheduling of lecturers and classes gives rise to a staggering number of combinations. To enumerate all the possible combinations or to make entries at random and starting from the beginning, in case of failure, leads to a prohibitively large number of trials. The algorithmic approach is also not practicable in view of the large number of quantified constraints and the selection of an objective function, which is really objective. A heuristic method comprising (a) an organization of data, (b) a pre-assignment distribution, and (c) assignment heuristics, is described. By the method, it is economically feasible to construct quality timetables, in an efficient manner, for technical institutions.

TRADITIONALLY, the timetable of an institution has been a vehicle for the administration to express much of its philosophy of education. The timetable gives practical shape to the academic policies concerning the subjects to be included in a course, the relative weights of subjects, the amount of laboratory work, and the allocation of time for tutorials, seminars, and discussion groups. Quality timetables are also required to provide for individualized instruction and to cater to the varied needs of students by providing both a large number of subjects to choose from and individual projects satisfying the varied interests, abilities, and aptitudes of students.

The timetable in a technical university has some special features, e.g., students spend quite a considerable amount of time in laboratories, workshops, and drawing halls spread over 2 to 4 consecutive hours. The large amount of time creates special problems which are not present in ordinary school or university timetables. Moreover physical facilities such as drawing halls, laboratories, and workshops with their machinery and equipment are much more expensive than physical facilities required for general education. As such, more effective utilization of these expensive resources is a *sine qua non* of good educational administration.

Problem

For any university, some of the data are fixed: (a) the number of working days in a complete cycle of a week, (b) the number of periods per day, (c) the number of classes and parallel sections, if any, and (d) lunch hour(s), which may be staggered in view of the capacity of the lunchroom or with a view to increasing utilization of resources.

The rest of the data may be based on the decisions of the academic bodies or heads of departments: (a) subjects for each course; (b) number of lectures—practical hours for each subject; (c) allocation of lecturer(s) for each subject; (d) special requirements such as two or more consecutive periods, "sets," or combinations with other classes, meeting at a particular hour, team teaching within a subject, any particular spread of instruction over the weekly sequencing of lectures, tutorials, and practicals, and maximum teaching loads and other resources maxima; and (e) non-availability of lecturers on certain hours due to personal or official reasons.

Moreover, it is desirable to spread the lectures in a particular subject over the entire week. The even distribution of lectures results in more efficient learning and helps in leveling resources. For example, a subject with three lectures a week may best be assigned on Monday, Wednesday, and Friday or Tuesday, Thursday, and Saturday. However, with limited resources, Monday, Tuesday, and Thursday or Tuesday, Wednesday, and Friday may also be acceptable.

The problem is to take into account all the above mentioned data, the human and physical resources of the university, the constraints imposed by special requirements including resources maxima, and provision of a time assignment for all classes and lecturers so as to obtain effective utilization of available resources of brain-power, expensive machinery, equipment, classrooms, and drawing halls while at the same time achieving a sound program—harmonious and well-knit in its component parts.

Construction of Timetable for a Technical University

Even for a medium-sized university, the scheduling of lectures and classes gives rise to a staggering number of combinations. To enumerate all the possible combinations in determining the 'optimum' solution is beyond the power of the most powerful computing machinery available. Almond's (1) suggestion for making entries at random in a blank timetable and starting from the beginning, in case of failure, will also lead to a prohibitively large number of trials for a real university.

The algorithmic approach also is not practicable in view of the astronomical number of combinations, large number of quantified constraints, and the selection of an objective function which is really objective. It is not surprising that Sherman (3), after trying to solve the scheduling problem for university classes by Integer Linear Programming wrote, "The . . . formulation, however, is not useful in treating problems of large size—say, when m is greater than 10 (where m is number of subjects) even for two set partitions." With his approximate stochastic method, he claims to deal with up to fifty Ss, whereas in a technical university, it is necessary to deal with hundreds.

Intelligent heuristics, based on experience, can provide economically feasible solutions by narrowing down the choice of combinations to a reasonable level. Heuristics accept qualitative specifications rather than quantitative ones, deal with local parameters, rather than global ones, and provide a flexible approach to alter parameters and thus achieve acceptable goals. A heuristic method follows which is given to construct quality timetables for technical institutes in a reasonable time.

Organization of Data

Every lecturer and class is given a code number. The requirements of "single," "double," and "triple" periods and "sets" are written down in standard form as shown in Table 1.

(a) If lecturer 3 is required to teach class 4 for two single and one double period per week, the requirement is recorded as shown against (a).

(b) If there is a "set" in which lecturers 4, 5, and 6 have to take care of classes 3, 4, and 7 for one single and two double periods, the lecturers and classes are entered in parentheses, as shown against (b).

(c) A combined period by lecturer 2 for classes 3 and 4 is entered as shown against (c).

(d) A 3-hour laboratory practice by class 4, when jointly supervised by lecturer 5 and assistant 11, is entered as shown against (d).

Appropriate remarks are made for the entries, in Table 1. The lecturer's availability on a particular day and hour, is indicated in a proforma by a "1," if he is available, otherwise by a "0." Another proforma is similarly prepared to record the availability of all the rooms, drawing halls, and laboratories for all the working hours of the university. A similar proforma records when the classes are free and when they are engaged. These proformas can be updated whenever a time assignment is made for any lecturer and class. Such an organization of data insures that all interrelated and interdependent facts, such as interfaculty lectures, are collected before starting the construction phase.

Table 1.—Organization of Data

	Lecturer(s)	Class(es)	S	D	T	Remarks
(i)	3	4	2	1	-	
(ii)	(4, 5, 6)	(3, 4, 7)	1	2	-	Set of 3 classes and 3 lectures
(iii)	2	(3, 4,)	1	-	-	
(iv)	(5, 11)	4	-	-	1	Lab. 'A'

Pre-assignment Distribution

The pre-assignment distribution enables decomposition of the large and complex problem to five or six small and relatively simple problems by spreading the teaching load over all the working days of the week. Since the laboratories and drawing periods generally take large amounts of consecutive hours and require supervision by two or more persons, and also because these practical activities are held in special rooms, the distribution of such periods is given top priority.

Then the subjects with a large number of periods per week will receive priority, e.g., five lectures per week for a 5-day working week will be distributed one period a day. The next priority is assigned to "sets."

Subjects with three periods per week will have the most desirable pattern as M, W, F or Tu, Th, Sa. Subjects with two periods per week can be spaced 2 or 3 days apart. While distributing the instructional load, it is desirable to keep a vigilant eye on the equitable distribution of individual lecturers and class loads. It is advisable to assign lecturers and classes on a day when they have the least number of periods already filled. Also the complexity of the problem can be leveled out by as nearly equal distribution of triple and double periods as possible. Last of all, the single periods are allocated by days so as to insure final adjustment of instructional load.

Assignment Heuristics

The daily instructional load, obtained by equitable pre-assignment distribution, is then time assigned by means of Assignment Heuristics. Priority is assigned in the following order: assignment of laboratory, drawing, and design periods; assignment of "sets"; assignment of double periods; and assignment of single periods. The possible assignments are found according to the periods lecturer i is available (P_{1_i}) and the periods class j is available (P_{c_j}). For example,

	Periods						
	1	2	3	4	5	6	7
P_{1_i}	0	1	0	1	0	0	1
P_{c_j}	1	0	0	1	0	1	1
$P_{1_i} \cdot P_{c_j}$	0	0	0	1	0	0	1

This indicates that lecturer i and class j both are available during the fourth and seventh periods only. Similarly, joint availability of the lecturer and the class is found out for double periods.

Assignment of laboratory, drawing, and design periods. Starting from a blank timetable, the maximum number of such requirements are adjusted in afternoon hours so as to make more morning hours available for lectures which require greater concentration. If a particular group of classes have "sets," then care should be taken that their laboratory periods and drawing periods come during the same hours so that more periods are available for joint requirements of "sets."

Assignment of Sets. "Sets" involving three lecturers and three classes or two lecturers and two classes can cause enormous problems unless dealt with at the early stages. The time assignment of "sets" requires the simultaneous availability of all the lecturers and classes involved. Generally, entries can be made in those hours which are already allocated for other classes. This results in the least interference with the opportunity for remaining requirements. If a "set" requires a single period only, it is assigned in an hour where only a single period can be assigned. That will save the available double periods for other entries.

Assignment of Double Periods. Similarly, entries are made for double periods during those hours when lecturers and classes are simultaneously available and the assignment results in least interference with the opportunity of remaining requirements.

Assignment of Single Periods. The remaining requirements of single periods are arranged classwise as illustrated in Table 2. The joint availability of lecturers and classes during seven periods of the day are indicated in the Boolean array by zeros and ones. P_{ij} indicates the total number of periods lecturer i is available to teach class j. The requirement with the minimum available periods is given priority. For time assignment, that period is selected which has the least number of ones in its column for that class. Such a selection causes least interference with the opportunity for remaining requirements. After making each time-assignment, all 1's (availabilities) in that line are converted into 0's. Similarly, in other class groups, all the 1's against that particular lecturer, during the assigned period, are converted into 0's. The binary pattern is similarly updated after each and every assignment. To satisfy remaining requirements, if any, some entries are interchanged with others. The above method is applied to construct a number of timetables in the university and school setups (2).

Table 2.—Assignment of Single Periods

Class	Requirement	Periods 1	2	3	4	5	6	7	P_{ij}
1	4,1	~~Φ~~	0	$\boxed{1}$	Φ	0	0	0	$\cancel{3}$ $\cancel{2}$
	6,1	0	$\boxed{1}$	Φ	0	0	0	0	$\cancel{2}$
	7,1	$\boxed{1}$	Φ	0	Φ	0	0	0	$\cancel{3}$ $\cancel{2}$
2	1,2	0	0	0	$\boxed{1}$	0	Φ	Φ	$\cancel{3}$ $\cancel{2}$
	6,2	0	0	$\boxed{1}$	0	Φ	Φ	Φ	$\cancel{4}$ $\cancel{3}$ $\cancel{2}$
	7,2	0	0	0	Φ	$\boxed{1}$	0	Φ	$\cancel{3}$ $\cancel{2}$

Summary

The timetable of a technical university gives expression to the philosophy of technical education and policies formulated by academic bodies of the university. It has the special feature of a considerable amount of practical laboratory, drawing, and design work spread over 2 to 4 consecutive hours. The timetable must plan and organize the efficient utilization of expensive laboratories, machinery, and drawing halls by resources leveling and by providing schedules which are conducive to effective learning.

Due to the complexity and size of the problem, the algorithmic approach has not been successful in giving a solution for real size institutions. For example, Sherman's Interger Linear Programming model carried him only as far as dealing with ten Ss. His approximate stochastic method enabled him to deal with up to fifty Ss. However, intelligent heuristics based on experience can provide economically feasible solutions.

The method, described in this paper, comprises a procedure for organization of data, a preassignment distribution, and assignment heuristics. The organization of data helps in the planning of interrelated and interdependent activities such as inter-faculty lectures. The pre-assignment distribution cuts the large and complex problem to a size which can be easily tackled. The distribution of instruction also leads to more efficient learning. Finally, the heuristics help in the time assignment of various categories of instructional load, in an economic manner, by assigning priorities to more critical activities such as multiple periods, "sets," and double periods.

FOOTNOTE

1. Postal address of the author is A. Wahab Khan, 43, A. Karim Rd., Lahore–5, Pakistan.

REFERENCES

1. Almond, M., "An Algorithm for Constructing University Timetables," *The Computer Journal*, 8: (no. 4) 331, 1966.

2. Khan, A. W., "Resources Allocation: Timetable in a Technical University," qualifying thesis, University of Birmingham, United Kingdom, 1967.

3. Sherman, G. R., "A Combinatorial Problem Arising from Scheduling of University Classes," *Journal of the Tennessee Academy of Science*, 38: (no. 3) 115, 1963.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Perception of Teacher–Educators in India About the Ideal Pupil

T. N. RAINA
Regional College of Education, Ajmer, India

M. K. RAINA
University College of Education, Indore, India

ABSTRACT

This study was designed to determine what concepts teacher-educators in India have of the ideal student in terms of what characteristics they believe should be encouraged and discouraged and to compare the results with the concepts of teachers in the United States. Torrance's Ideal Pupil Checklist was administered to one hundred teachers of education in teacher training colleges in Rajasthan, India. When the sixty-two characteristics of the Checklist were ranked, a rank-order coefficient of correlation of .76 was obtained between the ranks assigned by the Rajasthan teacher-educators and United States teachers. In general, the Rajasthan Ss emphasize the receptive nature of man and de-emphasize man's self acting nature more than United States teachers.

"THE TEACHER is the Brahma, he is the Vishnu, and he is also the Shiva. The teacher alone is the Absolute God, and so we bow to him." These oft-quoted lines identifying the teacher (Guru) with Divinity, represent the traditional respect extended to teachers in India, and Altekar (1), an authority on ancient Indian education, remarked that "the balance of evidence is in favor of the view that the teacher deserves the greatest reverence." The function of the teacher was to lead the scholar from the darkness of ignorance to the light of knowledge.

But times have changed. Teachers in India have been divested of divinity and reduced to the position of instructors, wage-earners, and technicians. The recent strikes and violence against teachers on the university, college, and school campuses, the burning of buildings and material property by the youth point to a severe conflict between tradition and modernity in India. But the teacher-technicians, according to observers, do not seem to have a realization of and an appreciation for the changes brought about by modern times. Their teaching methods are more in the traditional dictatorial attitude of "you must learn what I teach" rather than in the democratic technique of leading the students in free-thinking, rational judgment, and discovery for one-self.

The study described herein was designed to determine what concepts teacher-educators in India have of the ideal student in terms of what characteristics they believe should be encouraged and discouraged and to compare the results with the concepts of teachers in the United States.

Procedures

The Ideal Student Checklist. Torrance's (8, 9) checklist entitled "Your Ideal Pupil" was used in this study. The checklist consists of sixty-two characteristics arranged in alphabetical order. The instrument was originally developed by drawing upon empirical studies that differentiated some group or groups of highly creative people from a similar group of less creative people. The checklist was presented in English along with the Hindi translation. The translation was done by a board of professors, consisting of experts in Psychology, Hindi, and Education in the various teachers' colleges of the State of Rajasthan, India. The teacher-educators who knew English very well were requested to go by the English words when they felt any doubt about the Hindi equivalents. Information was gathered regarding the age, qualifications, experience in a teachers' college, and sex of the respondents.

Sample. The Ss were one hundred teacher-educators of the State of Rajasthan in each of the eleven teachers' colleges of this State. Of the eleven teachers' colleges, three are post-graduate, while one conducts, besides the degree program in education, a 4-year teacher education program. The other seven are graduate teachers' colleges in education. The Ideal Student Checklist was mailed to the staff members of the colleges in some cases, while in others it was administered by the investigators themselves. The checklists, that for some reason had to be discarded, numbered nine. The sample of teacher-educators constitutes nearly 70 percent of the total number of teachers in the eleven colleges of education. Eighty-four are male and sixteen female. The mean age of the sample of teacher-educators is 36.00 years with a standard deviation of 16.95 years. The educational qualifications of these teachers is, besides a master's degree in their respective disciplines like History, Economics, Hindi, Science, etc., a graduate or master's degree in Education. Their teaching experience extends from 1 to 15 years. The State of Rajasthan, before its independence in 1947, was composed of small and large princely states. The people in general are backward and poor—not an unusual phenomenon with the princely states and the desert areas. Some of the sample teacher-educators hail from various places in Northern India. The socio-economic status of the sample in general is a middle class one.

Analysis of Data. In the checklist, the respondents were requested to indicate by a single check the characteristics which they thought should be encouraged, by a double check the five characteristics deemed most important, and by a strikeout those characteristics which should be discouraged or punished. Each of the sixty-two characteristics in the checklist were given one point for a single check, two for a double check, and minus one for each strikeout. In this way, an index of desirability was obtained for each of the sixty-two characteristics. The characteristics were then ranked in terms of the values obtained, the characteristic with the highest value being ranked one, and so on.

Results

The results, after the sixty-two characteristics were weighted as indicated above, are presented in Table 1. This table presents the ranks of these characteristics based upon the responses of one hundred teacher-educators in the State of Rajasthan, along with the ranks obtained by Torrance from a sample of 1,512 United States teachers.

An examination of the table shows that the ten characteristics most liked by the teacher-educators of India are: (a) industrious, (b) considerate of others, (c) receptive to the ideas of others, (d) being obedient, (e) courteous, (f) does work on time, (g) well liked by peers, (h) sincere, (i) curious, and (j) independent in thinking. The ten characteristics that the teacher-educators liked the least are: (a) likes to work alone, (b) regresses occasionally, (c) non-conformist, (d) critical of others, (e) fault-finding, (f) a good guesser, (g) being timid, (h) bashful, (i) always asks questions, (j) disturbs class organization.

The rank-order coefficient of correlation between the ranks derived from the responses of teacher-educators in India and those derived from Torrance's sample of 1,512 United States teachers is .76. This reflects considerable similarity between the two groups but is lower than similiar correlations among different groups of teachers within India and among different groups of teachers within the United States (8). It is interesting to note, however, that compared with the United States teachers the Indian teacher-educators place more importance on affection, competitiveness, doing work on time, industriousness, obedience, popularity with peers, physical strength, quietness, receptiveness to the ideas of others, and willingness to accept the judgments of authorities. As a group, they assign less importance than the United States teachers to asking questions, being a self starter, determination, being energetic, independence in thinking, independence

Table 1.—Rank Order of Characteristics of the Ideal Student Checklist Based on the Responses of 100 Rajasthan Teacher-Educators and 1512 United States Teachers

Characteristic	Rank Order	
	India	U.S.A.
Adventurous	15	19
Affectionate	12	32
Altruistic	26	36
Always asking questions	54	38
Attempts difficult jobs	14	20
A self starter	21	12
A good guesser	57	53
Bashful	55	56
Becomes preoccupied with task	25	41
Considerate of others	2	1
Critical of others	59	46
Courageous in convictions	31	22
Courteous	5	8
Curious	9	4
Competitive	24	34
Desires to excel	27	21
Determination	40	6
Domineering	42	61
Disturbs class organization	53	60
Does work on time	6	13
Emotional	48.5	50
Emotionally sensitive	44	43
Energetic	37	14.5
Fault-finding	58	58
Haughty and self-satisfied	43	62
Healthy	22	7
Independent in judgment	29.5	16
Independent in thinking	10	2
Intuitive	23	30
Industrious	1	9
Likes to work alone	60	44
Never bored	35	40
Non-conforming	61	51
Negativistic	50	59
Obedient	4	25
Popular, well liked by peers	7	28
Persistent	41	23
Prefers complex tasks	33	39
Physically strong	13	37
Quiet	17	48
Receptive to ideas of others	3	11
Regresses occasionally (Playful, childish)	62	49
Reserved	52	47
Remembers well	32	24
Self-confident	16	10
Self-assertive	45	42
Self-sufficient	28	31
Sense of humor	14	3
Sense of beauty	29.5	17
Sincere	8	5
Spirited in disagreement	39	45
Strives for distant goals	34	26
Stubborn	46	57
Sophisticated	38	54
Timid	56	55
Thorough	20	18
Talkative	47	52
Unwilling to accept anything on mere say-so	48.5	35
Visionary	51	27
Versatility	18	14.5
Willing to take risks	36	29
Willing to accept judgments of authorities	11	33

in judgment, persistence, sence of humor, sense of beauty, foresightedness, and unwillingness to accept things without examining the evidence.

Discussion

Since its independence in 1947, India has been faced with tremendous problems in social, economic, cultural, and political fields. One of the urgent problems in national reconstruction is to educate students to help increase the national wealth and thereby to improve ultimately the general living standards of the people. In order to achieve this laudable aim the consensus among the Indian teacher-educators, voiced by the erstwhile Secretary and Adviser to the Ministry of Education, Government of India, Saiyidain (5), is that "it should be the business of schools to produce men and women who will be prepared to work throughout their life like good laborers and craftsmen, stretching themselves to the maximum extent in the service of good causes."

In the light of the above remark, it is easy to visualize why the teacher-educators prize industriousness as one of the most important hallmarks of the personality characteristics of an ideal student. The training college teachers themselves encourage those student-teachers who work for long hours (in the preparation of their lesson plans and the like), and those who win distinction in the university examination. Student-teachers who do not work 70 hours a week or more are looked down upon as idlers and as persons who should be assigned a low grade. Little do the teacher-educators realize that "many highly creative persons do not seem to be industrious because they spend some of their time sitting and thinking, not visibly busy" (7). The emphasis on industriousness is perhaps associated with the concern of the teacher's colleges to equip prospective teachers with all the "tricks-of-the-trade" during a short span of one session extending to 200 working days. Such a program consequently leaves little time for independent study, thinking, reflection, and understanding.

The sample teacher-educators placed being considerate of others, receptivity to the ideas of others, and being obedient, in the second, third, and fourth ranks. The creative person must be open to new ideas, but must examine everything before he accepts it as his own. This quality of intellectual honesty, sincerity, and independence is not compatible with a conformist society dominated by external examinations. This intellectual quality will bring him into conflict and make the creative child isolated.

The whole fabric of the culture of India, by and large, is designed to emulate and prize the virtues of obedience. The Taittiriya Upanishad and the Bhagvadgita have directly and indirectly laid down prescriptions for the devoted to follow rules and perform duty unquestioningly and without the expectation of a reward. Sampurnandand (6), the present Governor of the State of Rajasthan, emphasized consideration of others, the taking up of duties without insistence on rights, selfless service (Nishkam Karma), as some of the objectives of life and adds, "man alone of all creatures is capable of it." Therefore, it is not surprising that teacher-educators value obedience so highly.

There is a consistent demand by Indian educationists that education take heed of the racial and national culture and give the student an idea of what it means and what it has stood for throughout centuries. Luthra (3) is correct when she remarks that "the system of education in India is greatly authoritarian. This is the reflection of the authoritarian temper of Indian society, where difference in opinion with an elder is often regarded as disrespect for him. Teacher-educators as guardians and interpreters of the cultural heritage of this country would recommend the inculcation of the characteristic of courteousness because an open disagreement with those in authority would be looked down upon as an act of great discourtesy. But the difficulty is that the creative child may be too busy to be courteous—if not too busy with his hands, too busy with his mind. He may unwillingly disregard the shibboleths of a traditionalistic and conservative authoritarian society where he is required to touch the feet of elders, teachers, and others. Touching the feet of elders, whatever the age of the subject, is considered a symbolic representation of courtesy in this part of the country. Nevertheless, it is very essential to help the creative child to behave courteously to all, high and low, and in his own interest and in the interest of the society that needs him, perhaps, more than any other.

Teacher-educators considered the characteristic of "does work on time" important by assigning it the sixth rank in order of preference, and there is nothing surprising in this. Indian education, as Saiyidain remarked, has hitched its wagon not to the star, but to the lamp post, and educative efforts draw their inspiration and are guided by external examinations. Bloom's (2) observation that "examination is the sole basis for assigning marks and all rests on the single examination performance," is not the whole story. Teachers at all levels are asked to explain failures and in extreme cases it may involve the loss of one's job. Teacher-educators may be justified from a pragmatic consideration, but the implications of such an attitude from the point of view of creativity are grave indeed. Students who deviate from the routine and hide-bound conformity of doing things orderly, though their contributions may be unique and "germinal" in a sense, are often "rebellious." Irregularity and disorder are not necessarily bad and need not be looked upon with disfavor.

The personality characteristics discussed so far impose sharp restraints on the development of creative abilities. But against this dismal picture, it is encouraging to note that teacher-educators have assigned a high place to personality characteristics like sincerity, curiosity, and independent thinking. One of the characteristics of sincerity is fearlessness to pursue the truth as one sees it. Again, curiosity drives him to a ceaseless search of knowing more and more about himself and his environment. In his search he looks at the stars, his mind seeks to penetrate into the mystery of their being. The new does not hold him in awe but he wants to explore it, to grapple with it, and to understand it. But, if that is the definition of sincerity and curiosity, the institutional programs in the teachers' colleges in actual practice, do render the rankings given to those personality characteristics as suspect. Saiyidain's (4) observations are relevant:

> *They (training colleges) have tended to stress method and teaching devices and skills to such an extent that students get no chance for the play of their critical intelligence of aims and purposes and values, . . . at a training college, as a rule, freedom and self-expression are applauded in theory as they are scrupulously excluded in practice.*

The individual under a regimented social order has no need to trouble himself with independent thinking. But in a country like India, that has given to herself a republican constitution, the individual has to think for himself or else democracy and progress will not flower, or even come into existence. It is, therefore, encouraging that teacher-educators have assigned an important place to independent thinking. But is this so in actual practice? Independent thinking must be accompanied by independent judgment and the individual must stick to that judgment even though the majority may disagree with him. Being in an uncomfortable minority takes great courage.

How is man's great "human asset," creativity, to be protected and nourished? It is the belief of the investigators of this study that creativity and spontaneity cannot be directly produced and externally injected into an organism. But there is the possibility of bringing into existence a climate—sociological, psychological, and educational—into which creativity, originality, and spontaneity will blossom. The data yielded by this study suggest some of the things that must be done.

REFERENCES

1. Altekar, A. S., *Education in Ancient India,* The Indian Book Shop, Benares, 1934.
2. Bloom, B. S., "Some Observations on Examination in India," in *Evaluation in Higher Education,* University Grants Commission, New Delhi, India, 1961.
3. Luthra, G., "Concept of Divergency Reflected in the Imaginative Stories of Children in New Delhi, India," Minnesota Studies of Creative Behavior, Department of Educational Psychology, University of Minnesota, Minneapolis, 1966.
4. Saiyidain, K. G., *School of the Future,* The Indian Press, Ltd., Allahabad, 1934.
5. Saiyidain, K. G., *Education, Culture, and Social Order,* Asia Publishing House, Bombay, 1963.
6. Sampurnanand, All India radio broadcast, "The Future of Education in India" (mimeographed), Planning Commission, Government of India, New Delhi, 1954.
7. Torrance, E. P., "The Creative Personality and the Ideal Pupil," *Teachers College Record,* 65:220-226, 1963.
8. Torrance, E. P., *Rewarding Creative Behavior,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965.
9. Torrance, E. P., *Understanding the Fourth-Grade Slump in Creative Thinking,* final report on CRP 994, Bureau of Research, Office of Education Washington, D. C., 1967.

IN COMMUNITY life, we seek to achieve leadership and influence, we must master the art of getting along with everybody. We can express our ideas frankly, but we must give others the same right. We must *expect* to encounter opposition, lack of understanding and other annoying manifestations of human nature, for they are natural reactions. A wise man of our acquaintance spends little time with people whose opinions agree with his. He says, "What can I learn from them? I know what they know and they know what I know. So I seek out persons from whom I am likely to hear something different, something new, something challenging."—THE CURTIS COURIER

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Forecasting the Academic Achievement of Engineering Freshmen

LOUIS ROSS, University of Akron

ABSTRACT

An analysis of seven predictive measures was utilized to derive a regression equation for the prediction of student success in a specific academic discipline. The Wherry-Doolittle Method was employed to determine the best battery of predictive measures. Data were obtained on ninety-four freshmen. The regression equation was checked for its practical value as a predictive instrument for the student's entry, as well as continuation, in the field being pursued.

THE PURPOSE of this study was to assess the predictive value of various prognostic measures in forecasting scholastic achievement at the end of the first semester in the College of Engineering at the University of Akron. A corollary problem was to develop the regression equation for predictive purposes. The freshman engineering group numbered ninety-four.

The predictive measures available for analysis were: Average High School Mark; ACE Q-, L,- and Total Score; Cooperative English Reading Speed Comprehension Score; Bell Adjustment Inventory Scores; and the Pre-Engineering Ability Test Score (Form ZPA).

The criterion of academic success selected was the quality point-ratio earned at the end of the first semester by each student. It was defined as the number of quality points earned in each course multiplied by the number of hours credit allotted to the course, with the products summed and then divided by the total number of credit hours taken.

It was, of course, possible to include each appraisal measure in a battery of tests to be used in predicting the criterion. Such a procedure was neither sound nor desirable for a number of reasons. Aside from the laborious mechanical computations involved, some of the instruments could actually function as suppressor variables thus serving to decrease the utility of the battery for predictive purposes. Moreover, it is unlikely, when a large number of tests are available for the prediction of the same criterion, that more than three, four, or possibly five of them would be required to obtain an optimum predictive value.

The technique employed in selecting the best battery of tests for predicting a criterion is known as the Wherry-Doolittle Method (1). Briefly stated these are the principles of the method. The single test which has the highest r (zero-order coefficient of correlation) is deemed to be of highest value in the prediction of the criterion. The method then aids in the selection of the second test that will add the most to the prediction when combined with the first. The process is continued and a third test is selected which adds the most to the prediction when used in combination with the first two. If necessary, a fourth, fifth, etc., test is selected. After the selection of each test, a shrinkage formula is applied to calculate whether the corrected multiple coefficient of correlation (R) is larger than the previous multiple R. If the addition of a test results in no gain in R, the test is discarded, and no further tests are added. Thus, the minimum number of predictive measures giving the highest corrected multiple R is secured.

The rank order of predictive measures employing this technique was: Average High School Mark, ACE L-Score, and the Bell Emotional Adjustment Score, respectively. Gross scores were utilized in computing means, standard deviations, and product moment r's.

The following notation was employed in developing the regression equation: X_0 = the student's average first semester quality point-ratio; X_1 = the student's Average High School Mark; X_3 = the student's gross ACE L-score; and X_9 = the student's gross Emotional Adjustment score.
The means (M) and standard deviations (SD) were:

$M_0 = 2.06$ $M_1 = 83.72$ $M_3 = 62.44$ $M_9 = 6.62$
$SD_0 = 0.51$ $SD_1 = 4.89$ $SD_3 = 12.05$ $SD_9 = 4.78$

The six zero-order r's were:

$r_{01} = 0.511$ $r_{03} = 0.394$ $r_{09} = -0.006$
$r_{13} = 0.375$ $r_{19} = 0.318$ $r_{39} = 0.034$

The zero-order regression equations for estimating first semester quality point-ratios (a) from Average High School Mark alone, disregarding all other predictive measures; (b) from ACE

L-Score alone, disregarding all other predictive measures; (c) from Emotional Adjustment Score alone, disregarding all other predictive measures, were respectively:

(a) $X_0 = .053X_1 - 2.32$ $SE_{01} = .4384$
(b) $X_0 = .017X_3 + 1.00$ $SE_{03} = .4687$
(c) $X_0 = -.00064X_9 + 2.06$ $SE_{09} = .5100$
$SE_0 = .5100$

The standard error of estimate (SE) measures the reliability of the predicted score obtained from the regression equation. From this, it was evident that the quality point ratio earned by a student at the end of his first semester could be predicted with greater accuracy by taking into account only his Average High School Mark than if only his ACE L-Score was considered. The Emotional Adjustment Score, if considered alone, was useless as a predictor, since the value of the SE was the same as the SD of the Criterion. Hence, in the case of equation (c), the mean value 2.06 of the Criterion would always be taken as the best guess of a student's first semester quality point ratio regardless of his Emotional Adjustment Score.

It can be stated that approximately 68 percent of the Criterion scores predicted from the Average High School Mark only may be expected to differ from their actual values by not more than ± .44 points or from the ACE L-Score by only ± .47. The remaining 32 percent may be expected to differ from their actual values by more than ± .44 points, in the case of the Average High School Mark, and by more than ± .47 points in the case of the ACE L-Score.

The coefficient of determination (r^2) indicates the proportion of the variance of the Criterion score that is associated with each appraisal measure. These were: .261, .155, and .00 for Average High School Mark, ACE L-Score, and Emotional Adjustment Score, respectively.

To obtain the regression equation for predicting the first semester quality point ratio considering variation in only two of the three predictive measures, ignoring the effect of the third variable, certain first order partial r's were computed. The following results were obtained:

$r_{01.3} = .426$ $r_{01.9} = .537$ $r_{03.1} = .253$
$r_{03.9} = .394$ $r_{09.1} = .191$ $r_{09.3} = -.021$

The correlation coefficient $r_{01.3}$ measures the correlation between the Criterion and Average High School Mark for those students who have the same ACE L-Score, if the Emotional Adjustment Score is disregarded. The first partials for Criterion and Average High School Mark are only slightly different from the zero-order r of .511, whereas, the effect of partialling Average High School Mark is noticeable in the relationship between ACE L-Score and the Criterion ($r_{03.1}$ is .253, in contrast to r_{03} of .394).

The regression equations for predicting the Criterion (X_0) from the other predictive measures in pairs (if one omits the third variable) follow:

(d) $X_0 = .044X_1 + .0099X_3 - 2.25$ $SE_{0.13} = .424$
(e) $X_0 = .0115X_1 + .0185X_9 + 0.98$ $SE_{0.19} = .430$
(f) $X_0 = .00167X_3 - .000206X_9 + 1.02$ $SE_{0.39} = .469$

The reliability of the Criterion scores is estimated from equations (d), (e), and (f) is given by the standard error of estimate.

To obtain the regression equation for predicting the Criterion, taking into account all three of the predictive variables, Average High School Mark (X_1), ACE L-Score (X_3), and Emotional Adjustment Score (X_9), certain partial correlation coefficients of the second order have been computed. The following results were obtained:

$r_{01.39} = .4486$, $r_{03.19} = .2275$, and $r_{09.13} = .1544$,

where $r_{01.39}$ measures the correlation between the Criterion (X_0) and Average High School Mark for those students who have the same ACE L-Score and exhibit the same amount of Emotional Adjustment. The multiple regression equation to predict X_0 with these three scores combined was:

(g) $X_0 = .04966X_1 + .00890X_3 + .01468X_9 - 2.75$.

The reliability of a Criterion score estimated from equation (g) is given by the standard error of estimate, $SE_{0.139} = .419$.

The regression equation (d) may also be written in terms of standard scores as: $z_0 = B_{01.3}z_1 + B_{03.1}z_3$ where the B's are the beta coefficients. These beta coefficients indicate the relative importance of the Average High School Mark and the ACE L-Score in determining the value of the first semester quality point ratio. $B_{01.3} = .4226$ and $B_{03.1} = .2352$. Their sum is .6578. Hence, in the determination of the Criterion only by Average High School Mark and ACE L-Score, the former has 64.24 percent of the determining power and the latter has 35.76 percent.

The regression equation (g) in standard form is:

$z_0 = B_{01.39}z_1 + B_{09.13}z_9$ where $B_{01.39} = .47615$,
$B_{03.19} = .21028$, and $B_{09.13} = .13759$.

The sum of the beta coefficients is .82404. Hence, so far as the Criterion (X_0) is determined by Average High School Mark, ACE L-Score and Emotional Adjustment Score, the first has 57.78 percent of the determining power, the second 25.52 percent, and the third 16.70 percent.

The simple correlation coefficient r_{01} or $R_{0.1}$, provides a means of interpreting the linear regression equation when the Average High School Mark alone is used. Likewise, multiple correlation coefficients provide means of interpreting the regression equations when two or more independent predictive measures are employed. Thus, the multiple correlation coefficient R indicates the relationship between the Criterion and the predictive scores based on Average High School Mark and ACE L-Score, taken in combination. It had been previously shown that the corrected multiple R's were: $R_{0.1} = .5112$, $R_{0.13} = .5489$, and $R_{0.139} = .5575$.

Although $R_{0.139}$ is greater than $R_{0.13}$ which in turn is greater than $R_{0.1}$, the conclusion was reached, as a result of applying the F-test (variance ratio) of significance, that it was worthwhile to take into consideration a student's ACE L-Score, as well as his Average High School Mark in predicting his first semester quality point ratio, but it was not certain that a more reliable prediction was being made if his Emotional Adjustment Score was used only with his Average High School Mark or in combination with his Average High School Mark and ACE L-Score.

The index of forecasting efficiency indicates the percentage reduction in the errors of prediction for a given value of R. The values of E which denoted the index of forecasting efficiency were:

$R_{0.1} = .5112$ $E = 14.05\%$
$R_{0.13} = .5489$ $E = 16.31\%$
$R_{0.139} = .5575$ $E = 16.81\%$

Therefore, it may be stated that an R as high as .5489 reduces the error of prediction by less than 17 percent. However, since the values of R are each higher than .50 the corresponding regression equations are considered of some value for predictive purposes.

As a practical check of the value of the regression equation for predictive purposes, every tenth student starting with the fourth was selected from the Engineering group and his Average High School Mark and ACE L-Score were substituted in regression equation (d). The results are tabulated in Table 1. Of the predicted Criterion values, eight differ from the actual Criterion values by less than one standard error of estimate, i.e., by less than ± .42 points or less than half a letter grade. The remaining two values differ by less than two standard errors of estimate or less than a letter grade.

An additional application of the regression equation is found in the fact that for promotion to the pre-junior year in the College of Engineering the student must earn a minimum quality point ratio of 2.00 for all work taken and must complete at least 64 hours of work including all required general courses. It was then pertinent to inquire as to the chances an entering student might have of earning the required quality point ratio if his Average High School Mark and ACE L-Score were known at the time of admission.

If the assumption is made that there is a normal distribution of the predicted first semester quality point ratios and if all students are considered as having the mean ACE L-Score, then regression equation (d) is used to obtain the approximate results illustrated in Table 2. It is noted that only slightly more than 50 percent of the present group made the required quality point ratio of 2.00 (mean of the group is 2.06) at the end of the first semester.

Those students entering college with an Average High School Mark below the mean of 83.72 and only the average ACE L-Score had less than an even chance of earning a 2.00 first semester quality point ratio. The probability of achieving the 2.00 value declined rapidly with a decrease in Average High School Mark. Thus, with an Average High School Mark of 70, the entering student has about one chance in ten of achieving the required quality point ratio.

Table 1.—A Comparison of Actual Criterion Values With Predicted Criterion Values Obtained from Regression Equation (d)

Case	Average High School Mark X_1	ACE L-score X_3	Predicted Criterion X_p	Actual Crit. X_0	Difference X_0-X_p
7	79	63	1.75	1.2	-0.55
58	86	74	2.27	2.3	0.03
119	84	77	2.21	2.5	0.29
156	84	71	2.15	2.4	0.25
202	79	56	1.78	2.1	0.32
284	79	47	1.69	1.8	0.11
342	86	65	2.18	2.4	0.22
404	78	50	1.68	2.3	0.62
472	83	47	1.87	1.8	-0.07
518	86	78	2.31	2.4	0.09

Table 2.—Estimation of Probable Success of Engineering Students

Average High School Mark X_1	Most Probable First Semester Quality point Ratio X_p	Approximate Chance of Success
85	2.11	3 in 5
83.72	2.06	13 in 25
82	1.98	1 in 2
79	1.85	2 in 5
76	1.71	1 in 4
73	1.58	1 in 6
70	1.45	1 in 10

From Table 2 it can be seen that a student's ACE L-Score (as a measure of mental ability) would have to be above the mean ACE L-Score to compensate for a lower value than the mean of the High School Marks. Unless his Average High School Mark is at least 80 his chance of success in the Engineering program is very poor.

One additional point should be considered—the nature of the required freshman subjects. Students with unusual ability in mathematics or in physical sciences would have a good chance to raise the quality point average (based on 64 credits) since the sophomore engineering program requires more of these subjects and fewer general educational subjects.

REFERENCE

1. Stead, W. H.; Shartle, C. L., and others, *Occupational Counseling Techniques*, American Book Co., New York, 1940.

NEW BOOKS from the Publishers

Teaching

Dawson, George G. (ed.). *Economic Education Experiences of Enterprising Teachers* (New York: Joint Council on Economic Education, 1970), viii + 120 pp. Paper.

Duggan, E. P. and W. A. C. Stewart. *The Sociological Review Monograph 15: The Choice of Work Area of Teachers* (Staffordshire, England: University of Keele, 1970), iii + 149 pp. Paper $4.60.

Fey, James T. *Patterns of Verbal Communication in Mathematics Classes* (New York: Teachers College Press, 1970), xii + 92 pp. Cloth $6.25.

Johnson, Lois V. and Mary A. Bany. *Classroom Management: Theory and Skill Training* (New York: The Macmillan Company, 1970), ix + 453 pp. Paper.

Olivero, James L. and Edward G. Buffie. *Educational Manpower: Bold New Venture* (Bloomington, Indiana: Indiana University Press, 1970), xiv + 365 pp. Cloth $7.50.

Schmuck, Richard A. and Philip J. Runkel. *Organizational Training for a School Faculty* (Eugene, Oregon: Center for the Advanced Study of Educational Administration, 1970), xi + 191 pp. Paper.

College

American College Testing Program, Inc. *Quiz for College Presidents and all Others Who Care about Students* (Iowa City: American College Testing Program, Inc., 1970), 31 pp. Paper.

Current Issues

Alloway, David N. and Francesco Cordasco. *Minorities and the American City: A Sociological Primer for Educators* (New York: David McKay Company, Inc., 1970), x + 124 pp. Paper $1.95.

Goethals, George W. and Dennis S. Klos. *Experiencing Youth: First Person Accounts* (Boston: Little, Brown, and Company, 1970), xiii + 399 pp. Paper.

Howe, Harold and others. *Racism and American Education: A Dialogue and Agenda for Action* (New York: Harper and Row Publishers, 1970), xv + 164 pp. Cloth $5.95.

Urofsky, Melvin. *Why Teachers Strike: Teachers' Rights and Community Control* (New York: Doubleday and Company, Inc., 1970), x + 349 pp. Paper $1.95.

Children

Child Study Association of America. *Children's Books of the Year 1969* (New York: Child Study Association of America, 1970), 48 pp. Paper $1.50.

Child Study Association of America. *1970 Supplement to Recommended Reading about Children and Family Life* (New York: Child Study Association of America, 1970), 24 pp. Paper $1.

Goodman, Mary Ellen. *The Culture of Childhood: Child's-Eye Views of Society and Culture* (New York: Teachers College Press, 1970), xiii + 167 pp. Cloth $6.25. Paper $3.25.

Hill, John P. (ed.). *Minnesota Symposia on Child Psychology*, Volume 4 (Minneapolis: University of Minnesota Press, 1970), ix + 275 pp. Cloth $5.

Guidance

Cottingham, Harold F. (ed.). *Elementary School Guidance: Conceptual Beginnings and Initial Approaches* (Washington, D. C.: American Personnel and Guidance Association, 1970), 143 pp. Paper $3.

Research

Best, John W. *Research in Education* (second edition) (Englewood Cliffs NJ: Prentice-Hall, Inc., Publishers, 1970), x + 399 pp. Cloth $8.95.

DeVault, M. Vere (ed.). *Research, Development, and the Classroom Teacher: Producer/Consumer* (Washington, D. C.: Association for Childhood Education International, 1970), 79 pp. Paper $1.75.

Mouly, George J. *The Science of Educational Research* (New York: Van Nostrand Reinhold Company, 1970), xiii + 541 pp. Cloth $8.50. Student Workbook x + 457 pp. Paper $3.95.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Student Reaction to the Grade Contract

HUGH TAYLOR
University of Victoria, Victoria, British Columbia

ABSTRACT

The study describes the grade contract and assesses opinions of students who experienced the technique. Data were secured from seniors in an educational measurement and evaluation course. Analysis of student responses to a 14-item attitude scale showed a generally favorable opinion toward the grade contract system. A further comparison of a grade contract group with a control group showed that the former viewed the grade contract method as a fairer system of assigning grades than the conventional technique ($p < .01$). No statistically significant differences were found between the two groups on the validity of the course examination, the understanding of the objectives of the course, or the quality of the standards for the course.

OVER THE last 70 years, criticisms of grades have taken many different forms. Grades have been shown to be unreliable for individual teachers marking the same examinations and unrelated in meaning, within and across departments in the same university as well as across universities. The same criticisms, of course, hold today. It is difficult to find a student or a professor who is satisfied with the conventional grading system.

Grades have also been accused of causing anti-intellectual motivation and disabling anxiety in students. These factors are currently well documented by such strange bedfellows as the Students for a Democratic Society (1) and the Association for Supervision and Curriculum Development (2).

A number of minor adaptions have been made by universities in order to improve the communicability of their evaluative procedures. Some have used evaluations pointing to the range of students' strengths and needs in the mastery of fields of knowledge, powers of criticism, and philosophical maturity. Others are using a pass–fail system, particularly in a student's non-major field. Encouraging though these innovations may be, they still fall short of what conscientious educators deem desirable. One innovation in evaluation that seems to offer real hope for improved communication between students and teachers is the grade contract. Based on an underlying belief in individual freedom, the necessity of commitment, and personal responsibility, the grade contract operates on an individual, one-to-one relationship as other grading procedures do not.

The Problem

This article reports an attempt to assess student opinion regarding the use of the grade contract in a senior educational measurement and evaluation course. The psychological principles and philosophy that form the basis of the grade contract at the elementary and secondary school levels are clearly presented by Frymler (3:228–271). Shirts (4) has used the procedure and offers suggestions for its improvement at the university level. This report adds to the growing literature on the grade contract by focusing on student opinion regarding the new evaluative procedure.

The grade contract is a program of study and experience which is mutually agreed upon by the student and his professor. It is predicated upon the assumption that the objectives of the course have been clearly stated and ordered in terms of quantity and quality. An actual grade contract used by the author may be found in Appendix A. Two copies of the contract are prepared with the professor and student signing both and each keeping one for his own use.

It can be seen that the student chooses the grade he desires and then works out a series of agreements on how the grade is to be attained. If the student completes the agreement he receives his grade. However, if he meets only part of the agreement, only that level of the contract will be binding. Grades are cumulative in the contract. The quantity and quality of the work increase as the grade levels change from D to A. If the student does unsatisfactory work on any part of the contract he is allowed to try again. The tasks are judged on a pass–fail or satisfactory–unsatisfactory basis. Thus the emphasis is on learning and succeeding with a definite de-emphasis on testing and failing.

The grade contract is basically a democratic procedure. It attempts to promote maximum communication between student and professor, cooperative planning and action, and effective learning. The grade contract, therefore, should be tested in the educational marketplace and evaluated by consumers. What follows is a modest attempt to assess the grade contract in action.

Sample

The grade contract system was initiated in an educational measurement and evaluation course during the 1968 fall semester at Washington State University. The course is required for prospective teachers. Students attended two 1-hour large lecture sessions each week plus a 3-hour laboratory period. The author taught two lecture sections and a colleague taught a third section. Three teaching assistants conducted the laboratory sessions. A total of 240 students was enrolled. The sample consisted of twenty-three male and sixty-three female students chosen at random from three laboratory classes conducted by one teaching assistant. Of the eighty-six students in the sample, sixty-one were taught in the author's large lecture section. Only two students had previously participated in a modified grade contract system. As there was no systematic method used to place students in the laboratory sessions on the basis of achievement or other selective criteria it was assumed the students were representative of the total enrollment group.

Procedure

Two weeks prior to the conclusion of the course a 14-item Likert-type attitude scale developed by the author was administered. The instrument consisted of seven positively and seven negatively phrased statements relevant to the grade contract. For each statement there were six possible response categories expressing attitudes from strongly agree to strongly disagree with no neutral position. Students were required to respond to every statement. The items were scored so that a score of six indicated the most positive attitude. The maximum and minimum possible total scores were respectively 84 and 14.

Table 1.—Rank Order of Attitude Statements by Medians

RANK		MEDIAN
1	The grade contract projects are well spaced throughout the semester.	5.27
2	The grade contract does not allow for adequate communication between students and teachers about course objectives.	5.10
3	The requirements of the grade contract seem complicated and confusing.	5.09
4	The grade contract challenges a student without threatening him.	5.06
5	The standards for the grade contract are not high enough for a 400 level course.	5.04
6	The grade contract should be continued in future Education 401 classes.	4.94
7	The requirements of the grade contract allow students to work in areas of self interest.	4.80
8.5	The grade contract improves relations between students and instructors.	4.38
8.5	The requirements for the grade contract are relevant to class lectures.	4.38
10	The requirements of the grade contract allow a variety of educational experiences.	4.33
11	The grade contract does not help the student in more efficient learning for the course.	3.75
12	The grade contract simply represents another "gimmick" in education.	3.65
13	The requirements of the grade contract emphasize quantity rather than quality of work.	3.03
14	When the requirements of the grade contract are based on a Pass or Fail students are not motivated to work at optimum level.	2.80

During the last lecture of the course a questionnaire, developed by Washington State University, entitled Student Opinion of Teaching Effectiveness, was administered to a large group section. Responses from the fifty-seven students in attendance were obtained. Only four of the questionnaire items were relevant to the present study. They measured student opinion on the validity of examinations, the fairness of grading procedures, the clarity of course objectives, and the quality of the course standards. The responses of the students were compared with those of thirty-one students who had answered the same questionnaire at the conclusion of the previous semester. The latter students had not experienced the grade contract system. Students rated the opinion questions on a 12-point scale. Each rating was marked so that a low score of 1 indicated the most favorable opinion. Following is an example of the format and content of the questionnaire items.

FAIRNESS IN GRADING

A | | B | | C | | D | | E

Absolutely fair and impartial to all

Shows occasional favoritism

Constantly shows partiality

Results

Table 1 presents each attitude statement ranked according to median score. Table 2 compares the mean score of the two groups of students on the four opinion statements by use of the t-test.

Table 2.—t-test Comparison of the Means for the Grade Contract Group Versus the Control Group on Four Opinion Questions

Question	Mean Grade Contract Group	Mean Control Group	SD Grade Contract Group	SD Control Group	Significance	
					t	
Validity of Examinations	3.90	4.03	2.45	2.43	.23	NS
Fairness in Grading	2.25	1.63	1.21	.88	2.62	.01
Major Objectives Of the Course	3.25	2.82	1.66	1.96	1.02	NS
Standards	4.06	3.85	1.06	1.86	.56	NS

Discussion

The results reported in Table 1 show that the median rating of all but two of the fourteen attitude statements are above the midpoint of the scale. As such, one may conclude that the students who experienced the grade contract system expressed a relatively high positive attitude toward the procedure. It should be noted that with reference to the item ranked 13, which suggests students felt the grade contract emphasized quantity rather than quality, that the highest quality aspect of the course was demanded at the A level. A check showed that 30 percent of the students actually received a final grade of A. The low rating given to the item ranked 13, which means that students tended to slightly agree with the statement, therefore, might have resulted partly from the fact that 70 percent of the students had not experienced the highest quality aspects of the course. Further research should measure the interaction of student opinion and the grade level for which the student contracted.

The low rating of the item ranked 14 reported in Table 1, which means high student consensus with the statement, suggests that students are not motivated to work at an optimum level under the contract system. This reaction is one that might have been predicted. Students are conditioned throughout their school lives to expect that the highest grades should be given to those who obtain the highest scores on tests. In a great many courses, particularly those with large enrollments, test results form the sole basis for determining the letter grades students receive. With a background of experiences which emphasize comparative achievement testing, it is reasonable to expect students to feel a lack of motivation when written tests are de-emphasized and the passing level is placed at 50 percent correct. As a check on the achievement level, a comparison was made between the grade contract group's mean score (29.81, N = 150) and the mean score (31.76, N = 86) of the 1968 summer session class that had written the same final examination. A test of the difference between the means resulted in a t = 2.6 which is the level required for significance at the .01 level with 234 degrees of freedom. Future research should check the validity of this finding and possibly focus on the problem of determining methods of maintaining high levels of test performance in groups experiencing the grade contract system.

Table 2 indicates that students who experienced the grade contract system rated the validity of the examinations, the clarity of the course objectives, and the standards required of students no differently than the control group. However, the grade contract students ranked the fairness of course grading significantly higher than the control group. The author feels this finding focuses on one of the important advantages of the grade contract system.

Conclusion

The foregoing results appear to indicate that student opinion, while not overwhelmingly positive toward the grade contract, is still not opposed to it. Rather, the results, particularly those concerned with the fairness of the grading procedures, demonstrate a willingness on the part of students to approve grade contract systems. New procedures almost inevitably generate resistance. The grade contract is no exception. The results reported here, however, offer a basis for some optimism regarding minimal resistance, at least on the part of the students sampled.

REFERENCES

1. Davidson, C., "University Reform Revisited," *Educational Record*, 47: (no. 1) 5–10, Winter 1967.

2. *Evaluation as Feedback and Guide*, Association for Supervision and Curriculum Development, 1967 Yearbook, National Education Association, Washington, 1967.

3. Frymler, Jack R., *The Nature of Educational Method*, Charles E. Merrill Books, Inc., Columbus, Ohio, 1965.

4. Shirts, Morris A., "The Grade Contract," *The Educational Forum*, 32: (no. 4) 456–458, May 1968.

(Appendix is on following page.)

APPENDIX A

GRADE CONTRACT

Name ______________________ Last First Educ.401 Evaluation of Learning. Spring 1969

Lab. Identification Number ______________ Lab. Instructor R.B., H.W., M.W.

Lecture Section 1 2 3 Laboratory Section Day M T W Th F Time AM PM

Major ______________________

I will do the following things for a grade of A B C D	To demonstrate that I have done these things I will
D. 1. Fill out contract in duplicate	1. Submit contract to my laboratory instructor by February 28.
2. Read *Evaluating Pupil Growth* by Ahmann and Glock	2. Pass examinations at the 50% level or receive a Normalized T-score greater than 37* March 10-11 Chapters 1-5 April 28-29 Chapters 6-10 Finals Week Mainly Chapters 11-16
3. a. Prepare an abstract and criticism of an article on evaluation in my area of specialization**	3. a. Submit the abstract and criticism to my laboratory instructor by March 7*
OR	OR
3. b. Prepare two 4-option multiple-choice test items from assigned chapters in Ahmann and Glock that measure intellectual processes above the knowledge level**	3. b. Submit the items to my laboratory instructor according to the following schedule: Chapters 1-5 by February 28 Chapters 6-10 by April 18 Chapters 11-16 by May 16
C. Complete the D category and	
4. Construct an evaluation instrument capable of measuring progress toward specific goals. This may be a small group (3) activity**	4. Submit the evaluation instrument and an accompanying written report to my laboratory instructor by March 14 and report orally on the instrument to my (our) laboratory class.*
5. Evaluate a standardized achievement test.**	5. Submit a written evaluation report to my laboratory instructor by May 16*
6. Evaluate a standardized ability test**	6. Submit a written evaluation report to my laboratory instructor by May 23*
B. Complete C and D categories and	
7. Understand and be able to calculate simple descriptive statistics and apply them to educational problems**	7. a. Hand in a first set of statistical problems to my laboratory instructor by April 4* and a second set by May 2* OR 7. b. Pass two statistical tests at the 75% level by May 2*
A. Complete the B, C and D categories and either (a) or (b) below	
8. a. Do an original research study on:** TITLE:	8. a. Submit a proposed plan of my study to my lecture instructor by April 4 and a written report of the research by May 16*
8. b. Do a study of studies on:** TITLE:	8. b. Submit a proposed plan of my study to my lecture instructor by April 4 and a written report of the study by May 16*

*Will be graded satisfactory or unsatisfactory. If unsatisfactory the task may be repeated once. Exception: final exam may *not* be repeated. If final exam is unsatisfactory, you will receive a letter grade, determined by your instructor, *less* than that for which you contracted.

**See your instructor for specific details.

Date Signed__________ Student__________________ Instructor__________________

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Self-Perceived Solutions to Adult Problems

RALPH C. DOBBS, University of Missouri

ABSTRACT

The data in this investigation were secured through structured interviews conducted in the homes of adults residing in urban areas. A total of one hundred adults were interviewed by the investigator. The number of interviews was divided equally between two communities and all respondents were randomly-selected heads of households.

All expressed problems were classified under the direction of a panel into four categories, namely (1) economic, (2) educational, (3) physical, and (4) health. The investigation was accomplished through a carefully designed interview schedule consisting of open-end questions which elicited self-perceived solutions to problems contained in the above categories.

The self-perceived solutions were classified into four categories, namely (1) improved educational program, (2) improved economic opportunities, (3) improved laws and zoning ordinances, and (4) improved health and medical care.

The urban areas involved in this study represent not only the non-declining community but also the community that is declining in use value.

IT SEEMS that much of the research in adult education today is either openly or inferentially criticized, not for its lack of ascertaining adult problems, but for its lack of suggesting solutions to the problems which adults express. A closer examination of the studies relating to this phase of adult education provides some enlightenment to the proposition that self-perceived solutions can be utilized in helping adults solve many of their own problems.

Green (3) has suggested five probable reasons why expressed solutions are not utilized in building adult programs. First, those adults engaged in manual labor were physically and mentally fatigued at the close of the day and often did not become involved in the process of critical thinking. Second, many adults felt inferior and did not want to reveal their ignorance by entering a formal adult class. Third, adults often felt that a language difference existed between them and their leaders. Fourth, educational offerings were usually too abstract to be relevant to their daily experiences. Fifth, the adults were not usually approached by educators in the learners' own social environment.

Bryson (2) sounded a warning that educators should not place too much emphasis on "telling" adults but rather guide them into experiences through which learners may discover what they need to know to solve many of their own problems.

Lindeman (4) stressed the importance of understanding the desires of adults in order to build a program of participation in educational affairs. He implied that educators need to comprehend more clearly the factors operating in the lives of adults in order to develop meaningful learning experiences.

Bergevin and McKinley (1) maintained that active participation in educational programs is apt to result whenever adults are involved in all aspects of program development.

Purpose and Procedures

The purpose of the study can be divided into four major parts.

1. *Find the self-perceived problems of randomly selected adults in a declining community and in a non-declining community.*
2. *Find the self-perceived ways to solve the problems expressed by the respondents in the study.*
3. *Make comparisons of the self-perceived problems and solutions of the adults in the study.*
4. *Search for ways in which self-perceived solutions can be utilized in formulating an adult education program.*

This study involved a detailed, personal structured interview with a random sample of 5 percent of heads of households of a declining urban community and compared the problems and solutions to those of their comparable counterparts in a non-declining urban community. A declining community is defined as an area which is losing use value to such an extent that its residents are having difficulty returning economic rents, paying assessed taxes, and providing for fundamental repairs.

In order to secure authenticity as well as replicability regarding the methodology involved in this research investigation, a number of authorities were contacted for counsel. From these authorities, four were secured as panel members.

The members[1] carefully examined the interview schedule and made extensive suggestions as to how the instrument could be effectively employed. After pretests in pilot studies, the instrument was adjusted and the categories for analysis of data were formulated. Classification of the data was made by the researcher and the members of the panel.

The boundaries established were as follows:

> *Economic: Included all responses interpreted as problems relating to finances as well as to occupational difficulties.*
> *Educational: Included all responses interpreted as problems concerning schooling, training, learning, and understanding.*
> *Physical: Included all responses interpreted as problems of the adult involving shopping accommodations, building, and community facilities.*
> *Health: Included all responses interpreted as problems relating to health as the most dominant problem in the adult's life.*

All one hundred respondents were heads of households and were selected at random. All dwellings were listed consecutively beginning at the lot occupying the extreme northeast corner of the designated community boundary line. Consecutive numberings were followed throughout each community until all dwellings were numbered. By utilizing a table of random numbers, the needed number of households was secured.

All data were classified according to the criteria set forth by the panel and respectively recorded on tally sheets by the researcher and panel members.

In order to make meaningful comparisons of the responses of the interviewees, as shown to be advisable by pilot studies, it was necessary to secure a dominant self-perceived problem and a dominant self-perceived solution from each respondent. This particular approach yielded data of sufficient nature to be treated by chi-square statistical analysis.

The interview questions were personalized and preplanned. For example, "What is there about you or your community that makes it hard for you to do what you want to do?" and "What would help you remove or get around the things that stand in your path?" are typical questions included in the interview. All interviews were made in the homes of the individual respondents, and the expressed solutions were received in a face-to-face situation.

Table 1.—Percentage Distribution of Self-perceived Problems of Adults in a Declining and Non-declining Community

Self-perceived Problems	Type of Community Declining (N=50)	Non-declining (N=50)
Educational	34	32
Physical	18	42
Economic	34	16
Health	14	10

Since the panel of judges approved the criteria of classification of all problems as educational, physical, economic, or health, it seemed important to show in Table 1 the percentage distribution of the self-perceived problems in the two types of communities.

The economic type of problem represented slightly over one-third of the total problems found in a declining community. The largest single type of problem in the non-declining community was of a physical nature. Based on these data it seems that the two communities view physical and economic problems differently but health and educational problems similarly.

Table 2.—Frequency Distribution by Type of Problem Perceived by Adults in the Declining and the Non-declining Community

Type of Problem	Type of Community Declining (N=50)	Non-declining (N=50)	Total
Educational			
Actual	17	16	33
Expected	16.5	16.5	33.0
Physical			
Actual	9	21	30
Expected	15.0	15.0	30.0
Economic			
Actual	17	8	25
Expected	12.5	12.5	25.0
Health			
Actual	7	5	12
Expected	6.0	6.0	12.0

Further Treatment of Data

An effort was made to treat statistically the differences between the self-perceived problems of adults in the declining community and the self-perceived problems of adults in the non-declining community.

Table 2 shows the frequency distribution of the problems as perceived by the adults in both the declining and the non-declining community. These data are shown in order to reveal the *actual* cases involved and at the same time yield a statistical measurement as a further means of comparing the types of problems perceived.

Calculations revealed that the total chi-square value was 8.70 with three degrees of freedom. Chi-square tables from Wert, Neidt, and Ahmann (5) showed that a value of 7.81 was required to be considered significant at the .05 level.

Having established a significant difference between the self-perceived problems of adults in the declining community and the non-declining community, it seemed important to examine the self-perceived ways to solve their respective problems. Using the same categorization as set forth earlier, Table 3 was developed to make the comparison. A chi-square statistical analysis was made of the data which showed a total value of 7.04 with three degrees of freedom. The greatest difference occurred in the educational category, from which almost one-third of the total chi-square value was derived.

Since the overall purpose of this investigation is to clarify and improve the methods of adult learning, a further analysis was made on frequency distribution within each category. Hence the development of Table 4 seemed necessary. Table 4 shows the self-perceived ways to solve adult problems in a *declining* community. It seems rather important to an educational researcher who is concerned with declining communities to be able to conclude that 44 percent of the expressed ways to solve problems were educational.

Table 3.—Percentage Distribution of Self-perceived Ways to Solve Adult Problems

Self-perceived Ways to Solve Problems	Type of Community Declining (N=50)	Non-declining (N=50)
Improved Educational Program	44	26
Improved Economic Opportunities	22	38
Improved Laws and Zoning Ordinances	20	30
Improved Health and Medical Care	14	6

Table 4.—Frequency Distribution of Self-perceived Ways to Solve Problems of Adults in a *Declining* Community

Ways to Solve Problems	Frequency of Mention (N=50)
Improved Educational Program	
More night classes	7
More or better job training	6
Leadership training	5
Community Council	4
Improved Economic Opportunities	
Develop good paying jobs	5
Become home owners	3
Job location service	3
Improved Laws and Zoning Ordinances	
Property prohibition through zoning laws	5
More and better law enforcement	3
Equal rights regardless of color	2
Improved Health and Medical Care	
Improved medical program	5
Care for handicapped	2
TOTAL	50

Table 5 presents the same type of responses from adults who live in a *non-declining* community.

It seems relevant to note that in a non-declining community a smaller number of adults perceived fewer educational ways to resolve their problems than economic ways.

Conclusions

Based on the analysis of the entire study it is concluded that:

1. The self-perceived physical and economic problems of randomly selected adults in a declining community differ from the self-perceived physical and economic problems of their counterparts residing in a non-declining community.
2. The self-perceived problems categorized as health and education respectively were viewed similarly by adults from the declining and non-declining community.

Table 5.—Frequency Distribution of Self-perceived Ways to Solve Problems of Adults in a *Non-declining* Community

Ways to Solve Problems	Frequency of Mention (N=50)
Improved Economic Opportunities	
Shopping facilities	10
Develop better paying jobs	9
Improved Laws and Zoning Ordinances	
Zoning for proper development	10
Law enforcement	3
Equal work rights	2
Improved Educational Program	
Improved program for leadership	7
More adult classes and courses	6
Improved Health and Medical Care	
Better medical care and treatment	3
TOTAL	50

3. In the declining community more than 40 percent of the expressed ways to solve problems were of an educational nature.
4. In the non-declining community the largest category under self-perceived ways to solve adult problems was "Improved Economic Opportunities."
5. A closer examination and further study may reveal that self-perceived solutions can offer substantial help in building and implementing effective adult education programs.

FOOTNOTE

1. The panel consisted of Dr. Richard Turner, Director of Educational Research, Indiana University; Dr. Harold Moore, College of Education, Southern Illinois University; Dr. Mason Atwood, Bureau of Adult Studies, Indiana University; and Richardson Wood, President of Program Research, Inc., 35 East 84th Street, New York, New York.

REFERENCES

1. Bergevin, P. E.; McKinley, John, *Design for Adult Education*, Seabury Press, Greenwich, Connecticut, 1959, p. 23.
2. Bryson, Lyman, *The Drive Toward Reason*, Harper and Brothers, New York, 1954, p. 86.
3. Green, Ernest, "Why This Apathy?" *Adult Education*, Fall 1959, p. 136.
4. Lindeman, E. C., *The Meaning of Adult Education*, New Republic, Inc., New York, 1926, p. 8.
5. Wert, James E.; Neidt, Charles O.; Ahmann, J. Stanley, *Statistical Methods in Educational and Psychological Research*, Appleton-Century-Crofts, Inc., New York, 1954, p. 423.

Forthcoming Articles

Crago, Priscilla H. and Sushila Singhal; University of New York, Albany
"SEX DIFFERENCES IN THE SCHOOL GAINS OF MIGRANT CHILDREN"

Goddard, Alton R.; East Texas State University, Commerce
"RECENCY OF SECONDARY SCHOOL MATHEMATICAL PREPARATION AS AN INFLUENCE ON FIRST-YEAR COLLEGE MATHEMATICS ACHIEVEMENT"

Poppen, William A. and Charles L. Thompson; University of Tennessee, Knoxville
"THE EFFECT OF GRADE CONTRACTS ON STUDENT PERFORMANCE"

Stiles, Lindley J.; Northwestern University, Evanston, Illinois
"STATE OF THE ART OF TEACHER EDUCATION"

Urbick, Thelma M.; Western Michigan University, Kalamazoo and Douglas R. Gross; Arizona State University, Tempe
"COUNSELOR EDUCATION: A PROFESSION ON THE MOVE"

Webster, William J.; System-Wide Evaluation, Dallas Independent School District, Texas
"AN EMPIRICAL COMPARISON OF SELECTED ENROLLMENT PROJECTION FORMULAS"

White, William F. and Bert O. Richmond, University of Georgia, Athens
"SOCIOMETRIC PREDICTORS OF THE SELF CONCEPT AMONG FIFTH AND SIXTH GRADE CHILDREN"

Williams, Robert; University of Dayton Graduate School, Ohio
"TESTING FOR NUMBER READINESS: APPLICATION OF THE PIAGETIAN THEORY OF THE CHILD'S DEVELOPMENT OF THE CONCEPT OF NUMBER"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

The Effects of Type of Examination Anticipated on Test Preparation and Performance[1]

A. RALPH HAKSTIAN, University of Alberta

ABSTRACT

This study investigated the effects of anticipation of either an objective, combined objective and essay, or essay test on students' study methods and actual performance on both objective and essay examinations. In the first of two experiments—conducted under ordinary classroom conditions—introductory education students were told to anticipate one of the three midterm examination formats. In the second experiment—conducted with language arts students—the preparation period was tightly controlled to permit precise, unobtrusive measures of study methods. In both experiments each analysis of study method variables and various performance criteria revealed no reliable differences between the three treatment groups, suggesting that kind of examination expected did not affect amount or type of preparation, or actual performance on either examination type.

THERE IS much empirical verification of the common-sense notion that tests tend to facilitate learning at all educational levels (5, 6, 7, 8, 14, 15). There is evidence that students at least report that they prepare differently for an objective as opposed to an essay examination. In studies by Terry (17), Douglass and Tallmadge (4), Meyer (10), and Silvey (12), college students reported on questionnaires that they attempted to master smaller units of information when anticipating an objective examination and more comprehensive units when preparing for an essay examination.

There has, in fact, long been a contention among educators and writers in the area of measurement that essay examinations can make unique pedagogical contributions by promoting more desirable or, at least, different study methods, and presumably greater mastery of content (1, 16). Little empirical evidence has been advanced, however, to support this position. The results of published studies have not been consistent. Meyer (9), using recognition, completion-type recall, and essay-type recall examination sets, found no discernible effect of examination set on recognition tests, but a facilitating effect on delayed completion tests for the completion-type recall set. Meyer's study, however, was not conducted in the usual learning-testing situation—the Ss learned the material and took the examination in a laboratory setting. Thus, lack of the usual motivation may have constituted a threat to external validity. Vallance (18) found no differences in performance on either objective or essay tests between students anticipating an essay test and those expecting an objective test. Vallance's study, however, permitted a week's time lapse between preparation and testing, which may have washed out any actual effects. Sax and Collet (11) found that Ss anticipating a multiple-choice test showed higher performance on this type of examination than those expecting a recall test, but that on a recall examination, there were no differences between groups. Subjects, however, took three midterm examinations of the format they were to expect on the final examination. It seems possible that differences in difficulty between the two test types as well as a certain test-wiseness regarding multiple-choice items may have accounted for the higher performance of the multiple-choice-trained group over the recall-trained group on the multiple-choice test.

The purpose of the present study was to compare the effects of anticipation of either an objective, essay, or combined objective and essay examination on study methods and performance on objective and essay examinations, on items from the knowledge, comprehension, application, and analysis levels (2), and on a retention test given 2 weeks later. The study was composed of two experiments. The first investigated the effects of type of test anticipated in a typical classroom setting. The second constituted an effort to investigate the same phenomenon with much tighter control over the preparation period and direct observation of study methods. The two experiments will be reported separately.

Method and Results of Experiment 1

Subjects. A class of thirty-six Ss—mainly sophomores and juniors—enrolled in an education

Table 1.—Summary Statistics for Subjects[a] in Experiment 1

	Group Means[b]				
	1 (O)	2 (O&E)	3 (E)	Total	S.D. Total
SAT: Verbal	523.14	505.00	526.14	519.47	65.71
SAT: Quantative	475.43	536.20	519.71	507.74	93.77
SAT: Total	1009.75	1026.83	1047.17	1027.21	135.61
Cumulative G.P.A. (A=4;...;F=0)	2.49	2.36	2.64	2.50	0.70
Age	22.17	23.50	22.00	22.59	4.03

[a] N=36; 12 males, 24 females.

[b] 1=objective; 2=combined objective and essay; 3=essay examination anticipated.

foundations course at the University of Colorado during the summer session, 1967, was used in the experiment. The original forty-eight Ss in the class had been randomly assigned to the three treatment groups, with sixteen in each group. Eleven Ss dropped the course, however, leaving group sizes of twelve, twelve, and thirteen. One S of the treatment group of size thirteen was then randomly dropped from the analysis to ensure equal cell frequencies. The investigator acted as graduate assistant for the course. Table 1 gives descriptive information about the sample.

Procedure. Ss were randomly divided into three treatment groups—anticipating an objective, combined objective and essay, or essay midterm examination. In the first lecture, Ss were given a handout informing each of the textbook and lecture materials for which he would be responsible on the midterm examination and the type of examination to expect, along with sample questions. One week before the midterm, Ss were reminded of the test and of the type of examination each could expect. On the day of the midterm, Ss completed the following:

1. A 14-item questionnaire. This was constructed to ascertain time spent in preparation, type of preparation, techniques (underlining, note-making, etc.) used, and units (facts, principles, applications) stressed.

2. A 3-question essay examination. This test was graded independently and anonymously by both the investigator and another doctoral candidate in education. The mean of the two scores was used in the analysis. Inter-rater reliability was .84.

3. A 52-item objective examination. This test included true–false, matching, and multiple-choice items. The KR–20 reliability was .71. The fifty-two items were made up of thirteen items each from the knowledge, comprehension, application, and analysis categories (Bloom's *Taxonomy* (2)). Scores on these items were also analyzed separately.

Two weeks later Ss were given an unannounced 28-item retention test including seven items from each of the knowledge, comprehension, application, and analysis categories.

Design. The study method variables were analyzed by a one-way analysis of variance, using the BMD01V computer program with the three treatment levels for type of examination anticipated. The test performance variables were analyzed by a one-way analysis of covariance (BMD05V). With Scholastic Achievement Test (SAT), total scores covaried ($.21 \leq r\text{'s} \leq .81$ with criteria) except in the analysis of essay test scores in which grade point average (GPA) was covaried because of greater correlation with the criterion ($r = .51$). A correlation analysis was also conducted (BMD03D).

Results

Correlations between the objective, essay, and retention tests and various organismic and study method variables are presented in Table 2. Of particular interest are the correlations between degree of stress on facts (score $1 = \text{low}; \ldots 3 = \text{high}$ from the questionnaire) and scores on the objective and retention tests. Even after the effects of SAT total scores have been partialled out a substantial ($p < .05$) negative relationship remains in each case. Also, degree of stress on applications correlates with total objective test scores, —.53, after SAT total scores have been partialled out. A positive relationship, however, is seen between stress on applications and essay test scores, after the partialling out of SAT total scores. A stress on ideas and principles, in preparation, appears to be the most successful set, showing positive correlations with all test scores, although these correlations are not significant after partialling out the effects of SAT total scores.

Table 2.—Correlations Between Total Scores on the Objective, Essay, and Retention Tests and Selected Organismic and Study Method Variables (N=36)

	Test		
	Objective	Essay	Retention
SAT: Total	.72**	.36*	.54**
G.P.A.	.35*	.51**	.35*
Age	.12	.34*	.30
Hours Spent Studying	-.50** (-.26)[a]	-.11	-.26 (-.01)[a]
Degree of Stress on Facts (1=low;...;3=high)	-.43** (-.33)*[a]	-.06	-.43 (-.34)*[a]
Degree of Stress on Ideas, Principles (1=low, etc.)	.25	.36* (.29)[a]	.34* (.23)[a]
Degree of Stress on Applications (1=low, etc.)	-.30 (-.53)**[a]	-.29 (.35)*[a]	-.21
Degree of Underlining (1=not used; ...;4=used extensively)	.00	.03	-.05
Degree of Making Notes (1=not used, etc.)	-.29 (.07)[a]	.00	-.18

* $p < .05$; ** $p < .01$. Both are two-tailed probabilities.

[a] Partial correlation coefficients with effects of SAT: Total scores partialled out.

Table 3.—Results of One-Way Analyses of Variance of Eleven Study Method Variables From Experiment One

Dependent Variable (df=2)[e]	MS_e(df=33)	Group Means[a] 1 (O)	2 (O&E)	3 (E)	F	p
[b]Time at which studying began	1.0354	2.83	2.25	2.75	1.1537	>.25
Hours spent studying	57.4924	12.17	10.33	13.08	0.4092	>.25
[c]Underlining and highlighting	0.5000	3.42	3.33	3.42	0.0556	>.25
[c]Making notes and summaries	0.8485	2.17	2.33	1.83	0.9167	>.25
[d]Stress on facts	0.4848	2.33	2.17	2.17	0.2292	>.25
[d]Stress on ideas	0.2500	2.67	2.67	2.58	0.1111	>.25
[d]Stress on applications	0.3258	2.08	1.92	2.42	2.3876	>.10
Times reading text	0.2980	1.75	1.42	1.33	1.9576	>.10
Times reading supplementary text	0.4798	1.42	1.33	1.75	1.2158	>.25
Times reading lecture notes	2.4949	2.67	2.33	2.50	0.1336	>.25
No. of pages at one sitting	1793.2424	52.83	54.17	46.83	0.1021	>.25

[a] 1=objective; 2=combined objective and essay; 3=essay examination anticipated.
[b] 1=yesterday;...;5=ten days ago.
[c] 1=not used;...;4=used extensively.
[d] 1=low;...;3=high.
[e] Treatment MS reproducible by MS_e X F.

Results of analyses of the study method variables—data which were gathered by means of the questionnaire—are reported in Table 3. From Table 3, it can be seen that—with respect to reported time spent studying, thoroughness of review, techniques employed, and type of orientation regarding the breakdown of material—there were no differences between students studying for an objective, and those studying for an essay, or combined examination.

Table 4.—Results of One-Way Analyses of Covariance[a] of Seven Examination Criteria From Experiment One

Dependent Variable (df=2)[d]	MS_e(df=32)	Adjusted Group Means[b] 1 (O)	2 (O&E)	3 (E)	F	p
Total objective score	23.3510	27.21	29.11	27.26	0.6022	>.25
Essay score	189.0731	45.31	52.56	43.38	1.4432	>.25
Knowledge items score	3.6554	8.73	8.93	8.43	0.2036	>.25
Comprehension items score	4.8994	6.08	6.09	6.08	0.0001	>.25
Application items score	2.5608	5.51	5.59	5.49	0.0133	>.25
Analysis items score	2.6182	6.89	8.51	7.27	3.2801	>.05
[c]Retention items score	6.2801	12.35	12.65	13.31	0.4464	>.25

[a] SAT: Total scores constituted the covariate, except in the essay score analysis, in which the covariate was G.P.A.
[b] 1=objective; 2=combined objective and essay; 3=essay examination anticipated.
[c] df for MS_e=31. One case was lost from this analysis.
[d] Treatment MS reproducible by MS_e X F.

Table 5.—Correlations of the Objective Test Scores, the Two Item Type Scores, and Essay Test Scores with Selected Organismic and Study Method Variables From Experiment Two (N=26)

	Total Obj. Test Score	Knowledge Items Score	Comprehension Items Score	Essay Test Score
Age	-.54**	-.48*	-.49*	-.56**
Time spent reading handout	.37	.39*	.26	.22
Extent of underlining, highlighting (number of lines)	.64**	.58**	.58**	.43*
Extent of making notes and summaries (number of words)	-.10	.03	-.22	.18

*p < .05; ** p < .01. Both are two-tailed probabilities.

The final analyses in this experiment were performed on results of the various achievement tests. The dependent variables were the objective test, the essay test, the retention test (parallel to the objective but given 2 weeks later), and subtests of items from the knowledge, comprehension, application, and analysis categories, according to Bloom's *Taxonomy*. Results of the analyses are presented in Table 4.

From the analyses of the examination criteria, it is seen that anticipation of one type of examination as opposed to another had no apparent effect on actual performance on the tests and item types included. The dependent variables shown in Table 4 were also analyzed using a 3 × 3 factorial design—stratifying with three levels of SAT total scores rather than covarying them—in order to determine whether the treatment variable, type of test anticipated, interacted with the levels of SAT total scores. The results of these analyses showed essentially no treatment × SAT interaction, and, moreover, treatment mean effects consistent with those reported from the analyses of covariance. Thus, it is felt that considerable confidence can be placed in the results as reported in Table 4.

The results of the analyses of both study method variables and examination criteria in experiment 1 indicate, without exception, no apparent advantage for one type of test over another in terms of preparation or actual performance. The question arises as to whether the absence of significant differences between groups on the study method and performance criteria is perhaps the result of a lack of power due to an N of thirty-six. The lack of consistency, however, between criteria of the group with the highest mean, and the fact that the ordering of the means in the performance

Table 6.—Results of One-Way Analyses of Variance of Six Study Method Variables From Experiment Two

		Group Means[a]				
Dependent Variable (df=2)[c]	MS_e(df=21)	1 (O)	2 (O&E)	3 (E)	F	p
Number of readings	1.1041	1.86	2.29	1.90	0.3688	>.25
Total time spent (minutes)	107.0932	55.71	48.71	51.70	0.8079	>.25
Degree of underlining (lines)	1426.7259	73.71	48.57	52.30	0.9353	>.25
Degree of making notes (words)	572.8626	28.00	8.57	7.40	1.7564	>.10
[b]Degree of stress on facts	0.4707	2.71	2.14	1.80	3.6618	<.05
[b]Degree of stress on ideas	0.6551	1.86	2.29	1.90	3.1111	>.05

[a] 1=objective; 2=combined objective and essay; 3=essay examination anticipated.

[b] 1=low;...;3=high.

[c] Treatment MS reproducible by MS_e x F.

analyses reflected no *a priori* hypotheses would seem to indicate that the nonsignificant differences obtained were, in fact, random.

Method and Results of Experiment 2

Subjects. A class of twenty-six senior-level language arts students attending the University of Colorado during the 1967 summer session was used for this experiment. Subjects were required to take part in the experiment, and of the twenty-eight Ss enrolled in the course, only two were absent from the experiment. They were randomly assigned to treatment groups after they arrived for the experimental session. Thus, no selection factors (differential or otherwise) were present.

Procedure. Subjects were randomly divided into the same three treatment groups as in experiment 1—anticipating an objective, essay, or combined examination. The experiment was conducted in a single afternoon. Ss were gathered together for a special session on linguistics, after having been informed by their instructor that they would be evaluated on their performance in the session, and that this evaluation would constitute part of their course grade. Each S was given a copy of an introductory chapter from a descriptive linguistics text, along with instructions as to what type of examination to expect and sample questions. Subjects were allowed unlimited time to read the material (13 single-spaced pages), and were allowed to do what they wished with the material with respect to underlining, highlighting, summarizing, etc. When Ss thought they had mastered the handout on linguistics, it was taken away, and their reading time was unobtrusively recorded. Subjects then completed the following:

1. *An 8-item questionnaire. This was used to provide information on number of readings of the material, degree of stress on facts, and degree of stress on ideas. These data along with the directly obtained information on time spent, and amount of underlining, note-making, etc., constituted the study method variables investigated in experiment 2.*
2. *A 2-question essay examination with no time limit.*
3. *A 30-item multiple-choice objective examination with no time limit. This test had a KR-20 reliability of .69. The thirty items included fifteen items each of the knowledge and comprehension categories (Bloom's* Taxonomy), *scores on which were also analyzed separately.*

The handouts were examined, and the amount of underlining (number of lines) and note-making (number of words in notes and summaries) was recorded for each S.

Design. The study method variables were analyzed by one-way analyses of variance with the three treatment levels, using the BMD01V computer program. A 3 × 3 factorial design was used in the analyses of examination performance (BMD05V). The three treatment-groups—anticipation of an objective, essay, or combined test—were crossed with three levels of the age variable, which was found to correlate −.54 with the total objective test score, and −.56 with the essay score. Two Ss were dropped from the analyses of the study methods and performance variables because of missing data. All twenty-six were used, however, in a correlational analysis (BMD03D).

Results

Correlations between examination criteria and certain organismic and study method variables appear in Table 5. It can be seen from this table that age was strongly and negatively associated with performance. Also of interest are the high positive correlations between extent of underlining and highlighting of the material and performance on all criteria. This finding is not surprising, but did not appear in experiment 1. It must be remembered, of course, that in experiment 1 scores on this variable—degree of underlining and highlighting—were assigned from what Ss reported on a questionnaire. In experiment 2, inspection of the returned reading materials undoubtedly furnished more valid data on this variable.

Results of the analyses of the study method variables, data on three of which were gathered by direct observation (hopefully resulting in greater validity than student report), are presented in Table 6. Of the dependent variables reported in this table, only degree of stress on facts is seen to differ significantly between groups. Results of a multiple-comparisons analysis, using the Newman–Keuls procedure, showed that only groups 1 (objective) and 3 (essay) differed significantly ($p < .05$), with Ss expecting an objective test reporting a greater stress on facts. It must be emphasized that in light of the large number of analyses performed, the reader should not place very great confidence in this single significant difference. Since it is not corroborated by findings from experiment 1 and since group 1 does not show greater performance on factual level items in the test criteria analysis (Table 7), this difference may well be artifactual.

Overall, the results of the analyses of study method variables in experiment 2 are in close agreement with the corresponding analyses in experiment 1. The outstanding feature in both experiments is that the type of examination anticipated appears to make very little, if any, difference to time spent, breakdown of material, or techniques employed in preparation.

Results of the analyses of the examination criteria—scores on the total objective test, the knowledge and comprehension level items, and the essay test—are presented in Table 7. The results of the

Table 7.—Results of Two-Way Analyses of Variance[a] of Four Examination Criteria From Experiment Two

		F-Ratios for Effects[c]					
Dependent Variable	MS_e (df=15)	Treatment (df=2)	p	Age (df=2)	p	Treatment X Age (df=4)	p
Total objective test	18.3778	0.7059	>.25	2.6519	>.10	0.0652	>.25
Knowledge items	20.8722	0.5731	>.25	8.5441	<.01	0.2010	>.25
Comprehension items	5.4611	1.3735	>.25	2.9512	>.05	0.5493	>.25
Essay test	5.6611	0.2086	>.25	2.1345	>.10	0.4793	>.25

[a] A 3 x 3 factorial design with the three treatment levels crossed with three age levels was used.

[b] 1=objective; 2=combined objective and essay; 3=essay examination anticipated.

[c] Mean squares for effects reproducible by MS_e X F.

analyses of the examination criteria in experiment 2 are consistent with those of experiment 1—without exception, type of test anticipated had no effects upon actual performance on either examination type or items from different cognitive levels.

Conclusions

In summary, the evidence from both experiments first suggests that students do not prepare differently—in terms of time spent, organization of material, and techniques employed—whether they are expecting an objective, essay, or combined objective and essay examination. Secondly, it appears that regardless of type of examination anticipated, students will perform at the same level on either an objective or essay examination and on items requiring each of the various levels of cognitive functioning.

The use of various types of tests to foster various kinds of study and learning, although widely advocated (1, 3, 13, 16) would seem to be a practice based on intuitive appeal, but not convincingly supported by empirical research. In particular, the contention that a superiority of the essay examination is its ability to promote more desirable study methods and higher performance on tasks requiring organization, and deeper comprehension and analysis of information should be reevaluated in light of the evidence in the present study of no differences between groups in terms of study methods, the essay examination, or items from the higher levels of the cognitive domain.

FOOTNOTE

1. The costs of the production of this paper were partially met by U.S. Office of Education Grant #OEG-8-8-961860-4003(058).

REFERENCES

1. Adams, G. S., *Measurement and Evaluation in Education, Psychology, and Guidance*, Holt, Rinehart, and Winston, New York, 1965.
2. Bloom, B. S. (ed.), *Taxonomy of Educational Objectives: Handbook I Cognitive Domain*, David McKay Company, New York, 1956.
3. Cason, H., "The Essay Examination and the New Type Test," *School and Society*, 34: 413–418, 1931.
4. Douglass, H. R.; Tallmadge, M., "How University Students Prepare for New Types of Examination," *School and Society*, 39: 318–320, 1934.
5. Hertzberg, O. E.; Heilman, J. D.; Leuenberger, H. W., "The Value of Objective Tests as Teaching Devices in Educational Psychology," *Journal of Educational Psychology*, 23: 371–380, 1932.
6. Johnson, B. E., "The Effect of Written Examinations on Learning and on the Retention of Learning," *Journal of Experimental Education*, 7: 55–62, 1938.
7. Kellogg, W. N.; Payne, B., "The True–False Question as an Aid in Studying," *Journal of Educational Psychology*, 29: 581–589, 1938.
8. Keys, N., "The Influence on Learning and Retention of Weekly as Opposed to Monthly Tests," *Journal of Educational Psychology*, 25: 427–436, 1934.
9. Meyer, G., "An Experimental Study of the Old and New Types of Examination: I. The Effect of the Examination Set on Memory," *Journal of Educational Psychology*, 25: 641–661, 1934.
10. Meyer, G., "An Experimental Study of the Old and New Types of Examination: II. Methods of Study," *Journal of Educational Psychology*, 27: 30–40, 1936.
11. Sax, G.; Collet, L. S., "An Empirical Comparison of the Effects of Recall and Multiple-Choice Tests on Student Achievement," *Journal of Educational Measurement*, 5: 169–173, 1968.
12. Silvey, H. M., "Student Reaction to the Objective and Essay Test," *School and Society*, 73: 377–378, 1951.
13. Sims, V. M., "Objective Tests and Teachers' Measurements," *School and Society*, 36: 302–304, 1932.
14. Sones, A. M.; Stroud, J. B., "Review, with Special Reference to Temporal Position," *Journal of Educational Psychology*, 31: 665–676, 1940.
15. Spitzer, H. F., "Studies in Retention," *Journal of Educational Psychology*, 30: 641–656, 1939.
16. Stanley, J. C., *Measurement in Today's Schools*. (4th edition), Prentice–Hall, Englewood Cliffs, New Jersey, 1964.
17. Terry, P. W., "How Students Review for Objective and Essay Tests," *The Elementary School Journal*, 33: 592–603, 1933.
18. Vallance, T. R., "A Comparison of Essay and Objective Examinations as Learning Experiences," *Journal of Educational Research*, 41: 279–288, 1947.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

Factors Which Are Associated with the Principal's Behavior in Encouraging Teacher Participation in School Decision-Making

RICHARD A. GORTON

ABSTRACT

The purpose of the study was to investigate the relationships between several factors associated with the role of the principal in regard to teacher participation in school decision-making. A multivariate analysis of the data collected from a 10 percent random sample of high school principals in California indicated that the principal's personal role orientation and his perception of the expectations of his faculty and immediate superior were significantly related to his behavior in encouraging teacher participation in school decision-making.

IN RECENT years the issue of teacher participation in school decision-making has attracted considerable attention. Newspaper headlines have featured teacher demands for increased involvement in decisions influencing their professional welfare, and teacher organizations have become more militant in their attempt to obtain a more significant role for teachers in school decision-making.

A review of the educational literature on teacher expectations[1] for participation in school decision-making suggests that teachers prefer group leadership, i.e., leadership where teachers participate in making decisions about the school program, to principal-dominated leadership (1). Heretofore, however, research had not isolated the factors that are associated with the principal's efforts to expand or restrict teacher participation in the decision-making process. It was the identification of these factors and clarification of their relative contribution which constituted the focus of the investigation.

Related Research

Parsons and Shils (6), Sarbin (8), Getzels (3), and Gross (5) have developed paradigms of social behavior that include several variables pertinent to the problem. They postulated that an individual's social behavior in an institutional setting is affected by his own personal role orientation and by the expectations of the institution toward his role. An individual's personal role orientation is defined as his inclination to perform his role in a particular manner; institutional expectations are defined as those expectations for an individual's role behavior which are held by his relevant reference groups.

Studies examining the relationships of personal role orientation toward faculty participation, and to an individual's social behavior have been numerous (4, 7, 9). With the exception of Blau's exploratory study (2), however, these investigations focused on either personal role orientation variables or institutional expectations. They did not attempt to control the effects of one type of variable at the same time that they were examining the relationship of another variable. Blau noted, though, that studies aimed at explaining an individual's behavior as a function of his personal characteristics, without controlling for the effects of institutional expectations, do not isolate the effects of one kind of variable from the influence of the other (2).

Design

The investigation focused on two sets of variables which theorists have postulated would be associated with the behavior of the secondary school principal in facilitating efforts to participate in decision-making about the school program. The variables were (a) the principal's personal role orientation toward faculty participation, and (b) the institutional expectations for the principal's behavior, held by his relevant reference groups.

The two groups assumed to be most important in influencing the behavior of the principal on this issue were his faculty and superiors. Earlier studies had revealed that both groups possess expectations concerning the principal's role in encouraging faculty participation in decision-making, and each is related to the principal in the superior-subordinate hierarchy within the institutional setting. Since it is the principal's immediate superior who can most directly reward behavior, through favorable assessment of performance which is in accordance with his own expectations, and who can most directly apply sanctions for "incorrect" behavior, the study focuses on the expectations of the principal's immediate superior, rather than on his superiors as a group.

The study was conducted in one hundred high schools randomly selected from the California population of school districts with at least two high schools. All participants were sent the same questionnaire, which had been adapted from instrumentation developed at the Midwest Administrative Center for the purpose of studying decision-making in the schools.

Data on the principal's personal role orientation were obtained by asking each principal to respond to the questionnaire on the role that the principal should take in encouraging teacher participation in school decision-making. Information about the expectations of the principal's faculty and of his immediate superior for the principal's behavior was secured by requesting the principal to respond to the questionnaire as he perceived their expectations for his role. The investigator followed the conventional assumption that the principal is more likely to be influenced by what he perceives as the expectations of his faculty and immediate superior than by their actual expectations.

A 10 percent sample of each principal's faculty was randomly selected to describe the actual behavior (as represented in the questionnaire) of the principal in encouraging teacher participation in decision-making. A multivariate analysis was then employed to determine the relationships of each predictor variable with the behavior of the principal.

Findings and Conclusions

An analysis of the data revealed that significant relationships exist among the principal's behavior in encouraging teacher participation in school decision-making (r = .49), his perception of his immediate superior's expectations about teacher participation (r = .49), and his perception of his faculty's expectations about teacher participation in school decision-making (r = .43). All of the correlations were significant at the .01 level of confidence.

However, in order to assess the *independent* contribution of each predictor variable on the principal's behavior, a partial correlational analysis was performed which permitted the statistical control of possibly related predictor variables. The effects in each instance of using this statistical technique were noticeable and are presented in Table 1.

The use of statistical controls lowered the level of relationships between each of the predictor variables and the principal's behavior. However, the results presented in Table 1 indicate that the principal's personal role orientation about teacher participation in school decision-making is significantly associated with his behavior in encouraging teacher participation, independent of his perception of the expectations of his faculty and of his immediate superior.

Table 1.—Comparison of Pearson Product-Moment and Partial Correlations

Variables Correlated	PPM Correlation	Partial Correlation	Variable Statistically Controlled
1, 2	.49	.30	3
1, 2	.49	.25	4
1, 3	.43	.18	2
1, 4	.49	.25	2

Variables:
1: Teacher Ratings of Principal Behavior
2: Principal's Role Orientation
3: Principal's Perception of Faculty's Expectations
4: Principal's Perception of Superior's Expectations

Correlations at or above .23 are significant at .05 level.

Correlations at or above .30 are significant at .01 level.

The data presented in Table 1 also make it clear that the principal's perception of his immediate superior's expectations concerning his role in encouraging teacher participation in decision-making is significantly associated with his behavior with encouraging such participation, *independent* of his personal role orientation in that regard. However, a significant relationship was not found between the principal's perception of his faculty's expectations and his behavior when his personal role orientation was controlled.

Although it is tempting to conclude, based on these findings, that the principals in the study tended to be more sensitive to the expectations of their superiors than they were to the expectations of their faculties on the issue of principal encouragement of teacher participation in decision-making, a t-test did not reveal significant differences. But the fact that the partial correlations were the same (r = .25) for both the principal's personal role orientation and his perception of the expectations of his immediate superior when the effects of these two variables were alternately held constant (see Table 1) suggests that, as independent factors associated with his behavior in encouraging teacher participation, they exert equal influence.

Implications

The findings obtained from this study provide moderate support for the theory that the principal's behavior in an organizational setting will be influenced by his own personal role orientation, the expectations of his immediate superior and, to a limited extent, by the expectations of his faculty. However, perhaps more important, the study demonstrates the need for the use of statistical controls in investigations on variables associated with administrative behavior. Previous administrator research generally had not separated the effects of one kind of predictor variable from the influence of another in establishing a relationship with the dependent variable. This study indicates the clear need for such controls.

Some researchers may feel that the problem of identifying independent variables lies with using perceptions for the source of data on factors associated with the principal's behavior. However, a separate analysis conducted as a part of this study showed that the *actual* expectations of the principal's immediate superior were no more independent of the principal's personal role orientation than his perception of those expectations. It would appear that in either case—utilizing perceptions of or actual expectations—the problem of establishing an independent relationship between two variables must be solved through the use of statistical controls. Future studies on administrative behavior should take into account this methodological finding.

REFERENCES

1. *Administrator's Notebook,* 1: (no. 1), May 1952; 3: (no. 7), March 1955; 3: (no. 8), April 1955.
2. Blau, Peter M., "Structural Effects," *American Sociological Review,* 25:178–193, 1960.
3. Getzels, J. W.; Guba, E. G., "Social Behavior and Administrative Process," *The School Review,* 45:423–424, Winter 1957.
4. Gibb, Cecil A., "Leadership," *Handbook of Social Psychology, Lindzey, Gardner* (ed.), Addison-Wesley Publishing Company, Cambridge, 1954, p. 889.
5. Gross, Neal and others, *Exploration in Role Analysis: Studies of the School Superintendency Role,* John Wiley and Sons Inc., New York, 1958, pp. 11–70.
6. Parson, Talcott; Shils, Edward A., *Toward a General Theory of Action,* Harvard University Press, Cambridge, 1951, pp. 38–53.
7. Pierce, Truman M.; Merrill, E. C., Jr., "The Individual and Administrative Behavior," *Administrative Behavior in Education,* Campbell, Ronald F.; Gregg, Russell T. (eds.), Harper and Brothers, New York, 1957, p. 332.
8. Sarbin, Theodore R., "Role Theory," *Handbook of Social Psychology,* Volume II, Lindzey, Gardner (ed.), Addison-Wesley Publishing Company, Cambridge, 1954, pp. 223–258.
9. Stogdill, Ralph M., "Personal Factors Associated with Leadership: A Survey of the Literature," *Journal of Psychology,* 25:35–71, 1948.

FOOTNOTE

1. Rather than list individually the various studies on institutional expectations—a listing that would run for several pages—the reader is referred to the October issues of the *Review of Educational Research* for the years 1955, 1958, 1961, 1964, and 1967, for the bibliographical sources.

BOOK REVIEWS—Continued from page 296

indicates that "Precision Teaching is a powerful technique for changing behavior . . . and must be examined critically." The authors discuss the difference between education and indoctrination and the fact that indoctrination can occur no matter what type of teaching is utilized.

The authors anticipate that some readers will claim that the procedures "smack of bribery" and that children should work for the love of learning. The authors place the procedures in context by stating

> *There is nothing intrinsically evil about the systematic use of positive reinforcement as compared to natural social reinforcement. However, since for most people it is more desirable to use social rather than tangible reinforcers, we urge that whenever the tangibles are used that they be paired with the social. There are obviously some behaviors (e.g., work) that, under the present circumstances, need to be rewarded with a tangible outcome (money). However, even in these cases improved work efficiency would occur if the employer paired a pay check with social reinforcement (e.g., "You're doing a fine job Charlie.").*

Part II covers applications of the methods to children in normal classrooms, children with retarded behavior, socially deprived children, and children with severely deviant behavior. In these chapters many examples are given of how classroom teachers have used the methods outlined. The examples are realistic in that the difficulty of conducting an experiment in the classroom is noted and the teacher is cautioned that "many of those who attempt to use behavioral techniques for the first time try to do too much at once."

The authors address themselves to the fact that most teachers are interested in increasing the motivation of students and that motivation is thought of by the teacher as some sort of inner drive. The authors suggest that the only manifestation of the operation of motivation is overt behavior and urge that the index of motivation should be a behavioral rather than a verbal one. An example is given of the teacher who believes the child who says he is trying, despite all other behavior to the contrary. It has been the experience of the reviewer that it is difficult for teachers who have been taught to look for inner causes of behavior to accept other explanations. This combined with the tendency of teachers to place more credence in the spoken or written word than in demonstrated behavior has led many of them to reject application of learning theory principles despite vivid demonstrations of its success in their own classrooms. This is an additional reason to recommend that those who are interested in promoting the use of these principles in the classroom reinforce the teachers' awareness by such books as this one.

In the process of advocating the use of precision teaching, it is unfortunate that the authors felt it necessary to warn the readers against the use of intelligence tests stating "the very concept of measuring intelligence is highly misleading." The well publicized study of Rosenthal and Jacobson (1968) which indicated that when teachers are misinformed with respect to the IQ's of the children in their classes, the teachers and children tended to respond in terms of their expectations, is then quoted. This study is used as evidence that IQ results are meaningless although no indication is given of the flaws found by others in the statistical calculations of the Rosenthal and Jacobson study, which throws a different light on the interpretation of the results.

On the whole, *Changing Classroom Behavior* is a good book, and the material in each chapter is clear, concise, and to the point. The examples from research used to illustrate the principles are relevant, easy to understand, and convincing. Each chapter has a summary, which is followed by references and suggested readings. The book should be helpful to teachers and would make a good text for an in-service or workshop course.

Julia R. Vane
Professor of Psychology
Hofstra University

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

A Study of Personality Adjustment Among Adolescent Children with Working and Nonworking Mothers

DEANE D. NELSON,
Moorehead State College

ABSTRACT

This study sought to determine if there was a difference in personality adjustment as measured by the Minnesota Counseling Inventory (MCI) among 312 ninth-grade pupils who were categorized according to their mothers' work history. The principal statistical model used for analyzing the differences in groups was analysis of variance—treatments by levels design. The findings revealed that (1) in all instances, the personality adjustment (as evidenced by MCI scores in eight areas of adjustment) of boys was better when the mother worked full-time than if she had worked part-time or not at all; (2) the personality adjustment of girls did not follow a consistent pattern regarding the employment history of their mothers; (3) on the majority of the MCI scales, girls with nonworking mothers had better adjustment scores than girls with mothers who had worked either full-time or part-time.

MOST PARENTS face the question of whether the wife should work outside the home during the years that the children are attending school. Many families have no choice regarding the mother's working since the high cost of living makes it necessary for both parents to earn an income. Other families do, however, have a choice as to whether the mother should go to work. In many homes, the wife's income goes toward obtaining some of the "extra things" not normally available to the family due to financial limitations.

The purpose of this investigation was to determine if children with working mothers do in fact differ from children with nonworking mothers in personality adjustment as evidenced by MCI scale scores. More specifically, answers to the following questions were sought:

1. *Is there a difference in personality adjustment among children from homes where the mother works full-time, part-time or not at all?*
2. *Do boys and girls differ in personality adjustment among the maternal employment groups?*

Related Research

A relatively small amount of research has been done in the last 10 years regarding maternal employment and personality adjustment among children. Hand administered the California Test of Personality to 102 children in grades 4 through 8. Part of the group had working mothers and part had nonworking mothers. He found that the percentage of children that appeared to have adjustment problems was the same for both groups (4). Essig and Morgan gathered family life data on five hundred girls with working and nonworking mothers. It was found that girls whose mothers did not work were better adjusted in the area of family life and that there was a greater feeling of lack of love, understanding, and interest between many mothers and their daughters if the mother worked (2). Yarrow discovered that the incidence of overt rebellion and outbursts of protest was greater among children of working mothers (11). Roy maintained that the employment of the mother did not seem to affect the amount of affection in a family as perceived by children in homes where the mother worked (9). Glueck and Glueck discovered that emotional conflict was significantly more prevalent in delinquent boys whose mothers worked regularly or not at all (3). Maccoby felt that the fact that the mother worked had an unfavorable effect on the daughter's attitude toward her father. She felt that the father's role and status in the family was weakened by the mother's increased status due to her employment (6). Rouman reported that the majority of the children with working mothers lacked self-reliance, had more withdrawing tendencies, lacked social skills, and had antisocial tendencies (8).

Population and Sample

The parent population for this study was 1965–66 ninth-grade class from a large suburban school system near Minneapolis, Minnesota. The area of the community was very compact—approximately 2 miles wide and 5 miles long—and contained about 44,000 persons. There were around 12,000 pupils enrolled in K-12. The most common occupational group for males in the community was craftsman, foreman, and kindred workers, and for females it was clerical and kindred workers. The community had a median income of $7,700. The sample consisted of children in the school system who met the following criteria: (1) the child was living with both of his biological parents; (2) the child was Caucasian; (3) the child was a member of the 1965–66 ninth-grade class; (4) complete cumulative record information such as maternal employment history, and other significant data was available on the child; (5) the child's mother had a "pure" full-time or "pure" part-time work history or she had not worked during the life of

Table 1.—Number of Children Falling Into Each Maternal Employment Group According to Sampling Criteria Specifications

	Maternal employment groups		
	Full-time workers	Part-time workers	Housewife only
Sex	Number	Number	Number
Boys	52	60	157
Girls	52	77	146

N = 544

Table 2.—Number of Children in Each Maternal Employment Group According to Sampling Criteria Specifications

	Maternal employment groups		
	Full-time workers	Part-time workers	Housewife only
Sex	Number	Number	Number
Boys	52	52	52
Girls	52	52	52

N = 312, n = 52

the child. Table 1 shows the number of children in each maternal employment group that qualified for the study by these criteria.

Methodology

The MCI was administered to all members of the ninth-grade class. The MCI is an instrument used in junior high schools, secondary schools, colleges, and universities to acquire information regarding the personality characteristics of young people. Eight scores are obtained from this inventory. A validity or (V) score helps identify students who might be over-anxious to display socially acceptable characteristics. Other areas are family relationships (FR), social relationships (SR), emotional stability (ES), conformity (C), adjustment to reality (R), mood (M), and leadership (L). It should be noted that on all of the MCI scales, the lower the score, the better the adjustment in that particular area (1).

The students were also asked to complete a personal data questionnaire in order that certain information pertaining to their home background and maternal employment history might be obtained. A final sample of 312 children was randomly selected according to the previously outlined criteria (see Nelson (7) for detailed procedures regarding the investigation). Table 2 shows the number of children in each final maternal employment group. It should be noted that equal N's were used in the statistical analysis of data, in order that certain statistical tests could be utilized.

The principal statistical model used for analyzing the differences in groups was a Treatments × Levels design by Lindquist (5:121–155) using the test for "critical differences" in means when a significant F value was obtained (5:93–95). The primary purpose of the study was to determine if there was a difference in MCI scores among children from homes where the mother was employed full-time, part-time, or not at all. A secondary objective was to determine if boys and girls from different maternal employment groups scored differently on the MCI. Since boys and girls do tend to score differently on the MCI scales just as a function of sex, it was necessary to convert the raw scores for each scale into College Entrance Examination Board scores as outlined by Thorndike and Hagen (10:137–141). This conversion established the levels means for each scale at 500 with a standard deviation of 100. Thus the "built in" differences in the way that boys and girls tend to score on the MCI were eliminated. This conversion did not, however, change the existing differences in scores among the treatments nor did it affect any differences that existed among individual cell means. Therefore, it was possible to test each MCI scale for main treatment effect and interaction effect through analysis of variance. Lindquist referred to this type of raw score conversion as a "once corrected table" (5:110–112).

Results

The analysis of variance of validity, social relationships, conformity, reality, mood, and leadership scores respectively of children in the three groups showed no significant differences for either main effect or interaction on any of the MCI scales.

Table 3 shows the analysis of variance of family relationships scores of the children in the three groups. The F-value for groups was not significant. The F-value for interaction, however, was significant at the .05 level. Since the test for interaction was significant, it was permissible to test for critical differences between the mean family relationships scores of boys and girls for each maternal employment group. The "critical difference" test as outlined by Lindquist (5:93–95) may be used in testing for significant differences among means if all treatment groups are the same size, and thus, makes it unnecessary to compute individual t-tests for each set of means.

Table 3.—Analysis of Variance of Family Relationships Scores of Children with Working and Nonworking Mothers

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Maternal employment group	2	4,376.00	2,188.00
Sex	1	3.00	3.00
Interaction	2	72,629.00	36,314.50
Within	306	3,117,722.00	10,188.63
Total	311	3,194,730.00	

For groups, F (2,306) = .214
For interaction, F (2,306) = 3.564
For 2,306 degrees of freedom an F-value of 3.02 was required for significance at the .05 level.

Table 4.—Tests for Critical Difference Between the Mean Family Relationships Scores of Boys and Girls with Working and Nonworking Mothers

Maternal employment group	Mean (FR) scores Boys	Mean (FR) scores Girls	Difference
Full-time workers	475.25	516.90	41.65
Part-time workers	504.23	493.94	10.29
Housewife only	520.51	489.65	30.86

For 306 degrees of freedom a critical difference of 38.81 was required for significance at the .05 level. N = 312; n = 52 children per group.

Table 4 shows the tests for critical difference between mean FR scores of boys and girls in the three groups. The only significant difference was between boys and girls with full-time working mothers, with the boys being significantly better adjusted.

Table 5 shows the analysis of variance of emotional stability scores of the children. The F-value for groups was significant, but the test for interaction was not, however, significant at the .05 level. Since the test for groups was significant, it was permissible to test for critical differences between the mean ES scores of the children.

Table 5.—Analysis of Variance of Emotional Stability Scores of Children with Working and Nonworking Mothers

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Maternal employment group	2	73,883.00	36,941.50
Sex	1	3.00	3.00
Interaction	2	16,678.00	8,339.00
Within	306	3,120,637.00	10,198.16
Total	311	3,211,201.00	

For groups, F (2,306) = 3.622
For interaction, F (2,306) = .817
For 2,306 degrees of freedom an F-value of 3.02 was required for significance at the .05 level.

Table 6.—Tests for Critical Difference Between Mean Emotional Stability Scores of Children with Working and Nonworking Mothers

Mean (ES) Scores		Difference
Full-time Workers	Part-time Workers	
484.78	521.21	36.43
Full-time Workers	Housewife Only	
484.78	494.59	9.81
Part-time Workers	Housewife Only	
521.21	494.59	26.62

For 306 degrees of freedom a critical difference of 38.81 was required for significance at the .05 level. N = 312; n = 104 children per group.

Table 6 shows the tests for critical difference between the mean emotional stability scores of these children. It was found that none of the differences between the mean scores was significant at the .05 level. Therefore, the significant F-value for groups found in Table 5 was due to a nonsystematic cumulative effect somewhere in the table and not a difference between any two group means.

Summary

This investigation found that (1) there were no significant differences in MCI validity, social relationships, conformity, adjustment to reality, mood, and leadership scale scores for either sex or for the combined sexes among children whose mothers worked full-time, part-time, or not at all; (2) boys whose mothers worked full-time showed significantly better FR scale scores than girls whose mothers were employed full-time; (3) on all MCI scales, the scores of boys were better when the mother worked full-time than if she worked only part-time or not at all; (4) the personality adjustment of girls as shown by MCI scale scores did not follow any consistent pattern regarding the employment history of their mothers; (5) on the majority of the MCI scales, girls with nonworking mothers had better adjustment scores than girls with mothers who worked full-time or part-time.

REFERENCES

1. Berdie, R. F.; Layton, W. L., *Minnesota Counseling Inventory Manual,* The Psychological Corporation, New York, 1957.
2. Essig, M.; Morgan, D. H., "Adjustment of Adolescent Daughters of Employed Mothers to Family Life," *Journal of Educational Psychology,* 37:219-233, 1945.
3. Glueck, S.; Glueck, E., "Working Mothers and Delinquency," *Mental Hygiene,* 41: 327–352, 1957.
4. Hand, H. G., "Working Mothers and Maladjusted Children," *Journal of Educational Sociology,* 30: 245–246, 1957.
5. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education,* Houghton Mifflin Company, Boston, 1956.
6. Maccoby, E. E., "Children and Working Mothers," *Children* 5: 83–89, 1958.
7. Nelson, D. D., "A Study of School Achievement and Personality Adjustment Among Adolescent Children with Working and Nonworking Mothers," doctoral dissertation, University of South Dakota, Vermillion, 1968.
8. Rouman, J., "School Children's Problems as Related to Parental Factors," *Journal of Educational Research,* 50: 105–112, 1956.
9. Roy, P., "Adolescent Roles: Rural–Urban Differentials," *Marriage and Family Living,* 23: 340–349, 1961.
10. Thorndike, R. L.; Hagen, E., *Measurement and Evaluation in Psychology and Education,* John Wiley and Sons, Inc., New York, 1961.
11. Yarrow, M. R., "Maternal Employment and Child Rearing," *Children,* 8: 223–228, 1961.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 7, March 1971)

A Comparative Study of the Personality Structures of Persistent and Prospective Teachers

RICHARD R. DeBLASSIE, New Mexico State University

ABSTRACT

The extent to which personality differences exist between teachers and teachers-in-training was investigated. Thirty-two teachers (16 males and 16 females) and 32 persons with undergraduate liberal arts training (16 males and 16 females) who were undergoing teacher training at the post-baccalaureate level were administered Cattell's Sixteen Personality Factor Questionnaire. An analysis of variance indicated that the four subgroups were significantly different on four of the sixteen personality factors. The results suggest that only slight personality differences exist between teachers with an undergraduate teacher training background and prospective teachers with an undergraduate liberal arts background.

THE PERSONALITY of the teacher is a significant variable in the classroom, and indeed, some would argue that it is the most significant variable (5). Despite the importance of this variable and a great deal of research in this area, little is known about the nature of teacher personality, particularly in reference to differences between successful and non-successful teachers or between persistent and prospective teachers. There is also a paucity of research dealing with whether personality differences exist between teachers with an undergraduate teacher education background and those with an undergraduate liberal arts background who later trained to become teachers. The possibility that one sequence of experiences might be more conducive to successful practice than the other should be explored.

Although numerous studies of the relation between attitudes and teachers' observed characteristics have been done in discriminating between teachers' education in various types of institutions (liberal arts colleges, teachers colleges, etc.) (2, 3, 8), little research is available using personality inventories for this purpose. Several reasons might explain this. First, there are profound differences in what is meant by the term personality as measured by different personality inventories. This would make some researchers reluctant to use them for research since observations based on one definition will contradict observations based on another definition. A second issue hindering the research is the choice of instrument. The basic question in this regard is what devices are available for measuring personality, and from these how is one to choose the most appropriate measure? A third reason why there is a paucity of research is that the most frequent criterion of teacher effectiveness studies is the rating. Teachers and prospective teachers are rated by supervisors, principals, and the experimenters themselves in spite of the limitations of the rating method. Nevertheless, many experimenters feel more confident in using ratings as opposed to personality inventories in their research.

Examples of the type of research that has been done in this area is that by Lough (11) and Blum (1) who found that the Minnesota Multiphasic Personality Inventory (MMPI) was not a useful instrument for differentiating between student teachers and students in other curricula. LaBue (9), however, found that the MMPI did differentiate between students who showed a "persistent" interest in teaching (those students who completed a teacher-training program and accepted a teaching position) and those who did not. In addition, Hedlund (7) found that two MMPI scales (pa and L) differentiated between those who became good teachers and those who became poor teachers, and also those who failed to complete the program satisfactorily.

Some research utilizing Cattell's Sixteen Personality Factor Questionnaire (16 PF) has been done in this area (4, 6, 10, 12). However, the results seem contradictory and little research has been done with this instrument recently. Getzels and Jackson (5) suggest that the 16 PF has two specific advantages which warrant its use in this type of research: (1) it provides scores on factors that are not purely evaluative; and (2) it derives from an extensive program of both theoretical and empirical work carried on over a number of years.

The purpose of this study was to determine whether personality differences exist between persistent teachers with undergraduate training in teacher education and prospective teachers with an undergraduate background in the liberal arts. The extent to which personality differences exist between these groups was assessed by Cattell's 16 PF.

Procedure

For two samples of Ss, 16 PF profiles were obtained. The first consisted of a random sample of male and female Ss who had an undergraduate major in teacher education and who, in addition, had functioned as classroom teachers for at least 3 years. All of the Ss in this group were pursuing a Master of Arts Degree in Teaching (MAT) at the time of the experiment. This group will hereafter be referred to as the MAT group. The second sample consisted of a group of male and female Ss who had an undergraduate major in the liberal arts, and who had no classroom teaching experience, but were enrolled in a master's degree program designed to prepare them to teach disadvantaged youth. This program, known as the Teacher Corps Program (TC), was being carried on through a grant from the United States Office of Education. The two samples (MAT and TC) each consisted of thirty-two Ss, sixteen males and sixteen females for a total of sixty-four Ss.

Results

Analysis of variance was used to compare the personality structures of the four subgroups, while the t-test was used to test the significance of the differences between any two groups on each of the personality factors which indicated significant F ratios.

Table 1 includes the sten-score means, standard deviations, and F ratios on the 16 PF scales for each of the male and female subgroups, while Table 2 contains the various t-test combinations for those factors with significant F ratios. Table 1 indicates that the four subgroups were significantly different on four of the sixteen factors assessed by the 16 PF. Persistent teachers with undergraduate teacher training differ from prospective teachers with an undergraduate liberal arts background on the following factors: (1) Humble vs. Assertive; (2) Tough-minded; (3) Forthright vs. Shrewd; and (4) Placid vs. Apprehensive. More specifically, the following conclusions can be drawn from Table 2:

1. *MAT males are more humble, forthright, and tender-minded than MAT females.*
2. *MAT males are more humble, placid, and tender-minded than TC females.*
3. *MAT males are more placid than TC males.*

Table 1.—Means, SD's and F Ratios in Sten-score Units on the 16 PF Scales for the Four Subgroups

16 PF	MAT MALES X	SD	MAT FEMALES X	SD	T. C. MALES X	SD	T. C. FEMALES X	SD	F ratios
A Reserved vs Outgoing	2.75	1.19	2.63	.96	3.19	1.05	2.50	.96	1.31
B Less Intelligent vs More Intelligent	6.87	1.55	5.94	1.73	6.00	2.34	5.63	2.10	1.09
C Affected by Feelings vs Emotionally Stable	2.94	.73	3.19	.66	2.88	1.02	2.75	1.06	.67
E Humble vs Assertive	3.06	1.34	4.50	.89	3.13	1.15	4.38	1.02	7.80**
F Sober vs Happy-Go-Lucky	2.63	.81	2.50	1.10	2.56	.73	2.94	.97	.74
G Expedient vs Conscientious	2.94	.77	2.69	.93	2.69	.79	2.25	.93	1.66
H Shy vs Venturesome	4.31	.71	4.44	.89	4.06	1.13	3.75	1.12	1.55
I Tough-Minded vs Tender-Minded	3.94	1.23	2.69	.82	3.81	1.62	2.50	.89	6.44**
L Trusting vs Suspicious	2.50	1.21	2.94	.88	2.63	1.09	3.38	1.20	1.95
M Practical vs Imaginative	3.19	2.32	3.00	1.21	2.75	1.19	3.38	1.14	.65
N Forthright vs Shrewd	2.06	.77	3.13	1.04	2.31	.88	2.50	1.32	3.07*
O Placid vs Apprehensive	2.44	1.03	3.19	1.37	3.44	1.55	3.69	1.14	2.81*
Q_1 Conservative vs Experimenting	3.31	.79	3.75	1.24	3.44	.32	3.63	1.09	.48
Q_2 Group Dependent vs Self Sufficient	3.81	1.17	4.31	1.24	3.98	1.79	4.00	1.15	1.19
Q_3 Undisciplined Self-Conflict vs Controlled	3.25	1.07	3.56	.96	2.88	1.05	2.94	1.06	1.51
Q_4 Relaxed vs Tense	3.00	1.10	3.06	1.14	2.69	1.20	3.19	.84	.62

**=.01 Significance Level *=.05 Significance Level

Table 2.—Mean Comparisons of the Four Subgroups on Those 16 PF Scales with Significant F ratios

	MEANS				
FACTOR	MAT MALES	MAT FEMALES	Diff./M	"t"	P
E	3.06	4.50	-1.44	3.51	.01
I	3.94	2.69	1.25	3.38	.01
N	2.06	3.13	-1.07	3.06	.01
O	2.44	3.19	- .75	1.70	NS
	MAT MALES	T.C. MALES			
E	3.06	3.13	.07	.16	NS
I	3.94	3.81	.13	.26	NS
N	2.06	2.31	- .25	.83	NS
O	2.44	3.44	-1.00	2.13	.05
	MAT MALES	T.C. FEMALES			
E	3.06	4.38	-1.32	3.14	.01
I	3.94	2.50	1.44	3.69	.001
N	2.06	2.50	- .46	1.19	NS
O	2.44	3.69	-1.25	3.21	.01
	MAT FEMALES	T.C. MALES			
E	4.50	3.13	1.37	9.78	.001
I	2.69	3.81	-1.12	2.49	.05
N	3.13	2.31	.82	2.28	.05
O	3.19	3.44	- .25	.47	NS
	MAT FEMALES	T.C. FEMALES			
E	4.50	4.38	.12	.34	NS
I	2.69	2.50	.19	.63	NS
N	3.13	2.50	.63	1.43	NS
O	3.19	3.69	- .50	1.09	NS
	T.C. MALES	T.C. FEMALES			
E	3.13	4.38	-1.25	8.93	.001
I	3.81	2.50	1.31	2.85	.01
N	2.31	2.50	.27	.92	NS
O	3.44	3.69	- .25	.51	NS

4. *MAT females are less humble and tender-minded and more forthright than TC males.*
5. *MAT females and TC females are not significantly different.*
6. *TC males are more humble and tender-minded than TC females.*

Discussion

It is apparent from the results of this study that slight personality differences exist between persistent and prospective teachers. The results seem to agree with the MMPI studies of LaBue (9) and Hedlund (7) which indicated slight personality differences between students who showed a "persistent" interest in teaching and those who did not.

Several implications from this study might merit some consideration. The fact that persistent and prospective teachers do not seem drastically different in terms of personality structure suggests that regardless of the type of undergraduate education a teacher receives (teacher education vs. liberal arts), personality structure may not influence his success as a teacher (when success is defined as persistence). It should be borne in mind, however, that persistence is only *one* of many criteria that should be used in attempting to define the "successful" teacher.

The results and conclusions of this study should be cross-validated in order that the findings take on sounder meaning. The results of the study, at this point in time, should be considered only indicative of possible trends because of the small number of Ss used. The use of a larger N would insure a correction of the limitations imposed by a small number of Ss.

If replication and cross-validation of the present study using larger N's would indicate similar results, a logical next step in a series of studies would be that of attempting to determine (through correlational techniques, simple and multiple) the extent to which other variables, as well as personality structure, might be related to success as a teacher. The use of the analysis of variance in the present study should be considered as a first possible step in attempting to isolate variables which may influence success in teaching. The suggested correlational techniques would enable the researcher to actually determine the contributions that a number of variables, considered simultaneously, might make toward the variance in terms of the criterion variable (teaching success). It is hoped that the present research will stimulate much needed research in terms of success in teaching.

REFERENCES

1. Blum, L. P., "A Comparative Study of Students Preparing For Five Selected Professions Including Teaching," *The Journal of Experimental Education*, 16:31–65, 1947.
2. Callis, R., "Change in Teacher-pupil Attitudes Related to Training and Experience," *Educational and Psychological Measurement*, 10:718–727, 1950.
3. Cook, W. W.; Kearney, N. C.; Rocchio, P. D.; Thompson, A., "Significant Factors in Teachers' Classroom Attitudes," *Journal of Teacher Education*, 7:274–279, 1956.
4. Erickson, H. E., "A Factorial Study of Teaching Ability," *The Journal of Experimental Education*, 23:1–39, 1954.
5. Getzels, J. W.; Jackson, P. W., "The Teacher's Personality and Characteristics," in Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand McNally and Co., Chicago, 1963, pp. 506–582.
6. Hadley, S. T., "A Study of the Predictive Value of Several Variables to Student Teaching Success as Measured by Student Teaching Marks," *Teacher's College Bulletin, State Teachers College, Indiana, Pennsylvania*, 60:1–10, 1954.
7. Hedlund, P. A., *Cooperative Study to Predict Effectiveness in Secondary School Teaching: Third Progress Report*, University of State of New York and State Department, Albany, 1953.
8. Kearney, N. C.; Rocchio, P. D., "The Effect of Teacher Education on the Teacher's Attitude," *Journal of Educational Research*, 49:703–708, 1956.
9. LaBue, A. C., "Personality Traits and Persistence of Interest in Teaching as a Vocational Choice," *Journal of Applied Psychology*, 39:362–365, 1955.
10. Lamke, T. A., "Personality and Teaching Success," *The Journal of Experimental Education*, 20:217–259, 1951.
11. Lough, Orpha, "Women Students in Liberal Arts, Nursing, and Teacher Training Curricula and the Minnesota Multiphasic Personality Inventory," *Journal of Applied Psychology*, 31:437–445, 1947.
12. Montross, H. W., "Temperament and Teaching Success," *The Journal of Experimental Education*. 26:711–716, 1957.

FIELD NEWS

Willard E. North
Central Missouri State College

ONE OUT OF TEN OF THE largest school systems in the United States has community and/or parental control of at least one function in one or more elementary schools. This is the major finding of a survey conducted by the Center for Governmental Studies of Washington, D. C. Based on replies from 413 of the 545 school districts with 12,000 or more pupils, the Center's survey reports that forty-four have some degree of community control. After-school programs for children and other community uses of school facilities are the functions most frequently controlled. Other examples of parents/community control are in curriculum adoption and budget preparation (five districts) and control of discipline, site selection, and building design (four districts). No districts reported full control over selection of principals or teachers, although community boards in New York City and Washington, D. C., have partial control of personnel selection in sub-city school districts

Nearly five out of six districts indicate that they provide parent and/or community representatives an advisory role in one or more of ten functions. More than half provide this role for four or more functions, although in some instances the advisory role is for only a few schools in the system. The Center notes a tendency for the large school systems to provide a broader range of advisory roles for parents and other residents. As in community control, the Center found that after-school programs for children and other community uses of school facilities provide the most frequent advisory role for parents and other residents while school systems permit the least citizen advice in personnel selection. In between are building design, site selection, discipline, and budget preparation.

Using its present information as a background, the Center intends to embark upon an in-depth study of the experiences in a number of school systems with more than average community participation. This study is one of several which the Center has underway concerning administrative decentralization, citizen participation, and community control. Other current studies relate to state and local manpower programs, community corporations, and little city halls. For further information about any of these projects contact: Howard W. Hallman, President, Center for Governmental Studies, 1701 K Street, N.W., Suite 906, Washington, D.C. 20006.

* * * * * *

THE ESSO EDUCATION Foundation has announced grants for the academic year 1970–71 totaling $3.1 million, with $1.6 million allocated for support of experiments and innovations in higher education. This year's grants will go to some three hundred colleges, universities, and educational organizations through seven Foundation programs: SPUR (Support for Promoting the Utilization of Resources), Educational Research and Development, Presidential Contingency, Capital, Engineering in Science, Special Assistants, and Incentive.

The SPUR Program underwrites innovations aimed at improving the efficiency and effectiveness of undergraduate education in particular, while the Educational Research and Development program supports projects and experiments that have the potential to benefit either graduate or undergraduate education. In 1970–71, $1.6 million will be granted to at least two programs.

The $1.5 million being awarded under the Foundation's other programs reflects it continued interest in other phases of higher education. In 1970–71, 181 private institutions will receive Presidential Contingency grants totaling $505,000. The President of each institution will be asked to use the grants to underwrite the costs of unbudgeted items for activities that he believes will help further his institution's educational objectives.

Another $430,000 in capital grants will go to twenty-two private colleges and universities for development programs, specific building projects, and the improvement of laboratory, library, and other educational facilities. Sixty-eight engineering and science departments at forty-six public and private institutions will receive a total of $282,000 under the Engineering and Science Program. These unrestricted funds will be expended at the discretion of the department chairman. Twenty-three Special Assistance Grants totalling $170,750 will provide operating funds for specialized areas of higher education and organizations serving higher education. Lastly, $115,000 will be available under the Incentive Program for matching employee contributions to colleges and universities.

* * * * * *

TWO EDUCATIONAL CONFERENCES, one consisting of regional laboratories and the other made up of university-based research and development centers have merged into the Conference for Educational Development and Research. The two conferences voted to join to enable their member organizations to better cooperate in the efforts of providing educational innovations for the nation's schools.

The newly formed conference consists of twenty-five institutions, all funded in part by Title IV funds administered by the U.S. Office of Education. Fifteen of the institutions are regionally based educational laboratories and ten are university-affiliated research and development centers.

Edwin Hindsman, Executive Director of Southwest Educational Development Laboratory in Austin, Texas, was elected chairman of the new conference. Max G. Abbott, Director of the Center for Advanced Study of Educational Administration at the University of Oregon, was chosen chairman-elect. The conference's executive committee was formed with the election of John K. Hemphill, executive director of the Far West Laboratory for Educational Research and Development in Berkeley, California; Lawrence D. Fish, executive director of the Northwest Regional Educational Laboratory at Portland, Oregon; James W. Becker, executive director of Research for Better Schools, the Regional Laboratory in Philadelphia, Pennsylvania; John L. Holland, executive director of the Center for Social Organization of Schools at Johns Hopkins University, Baltimore, Maryland and Nate L. Gage, acting director of the Center for Research and Development in Teaching at Stanford University, Palo Alto, California.

Although the laboratories and centers have maintained separate conferences for several years, the institutions have cooperated in numerous joint efforts. Most recently the two conferences joined in funding an Information Office located in Denver, Colorado. The Information Office will continue to serve the laboratories and centers under the auspices of the new conference.

The fifteen regional laboratories in CEDR are Appalachia Educational Laboratory, Charleston, W. Va.; Center for Urban Education, New York, N.Y.; Central Midwestern Regional Laboratory, St. Ann, Mo.; Eastern Regional Institute for Education, Syracuse, N.Y.; Education Development Center, Newton, Mass.; Far West Laboratory for Educational Research and Development, Berkeley, Cal.; Mid-Continent Regional Educational Laboratory, Kansas City, Mo.; Northwest

Regional Educational Laboratory, Portland, Ore.; Regional Education Laboratory for the Carolinas and Virginia, Durham, N.C.; Research for Better Schools, Philadelphia, Pa.; Southeastern Education Laboratory, Atlanta, Ga.; Southwest Educational Development Laboratory, Austin, Texas; Southwest Regional Laboratory for Educational Research and Development, Englewood, Cal.; Southwestern Cooperative Educational Laboratory, Albuquerque, N.M.; and Upper Midwest Regional Educational Laboratory, Minneapolis, Minn.

The ten R & D centers include the Center for the Advanced Study of Educational Administration, University of Oregon, Eugene; Center for the Study of Evaluation, University of California, Los Angeles; Center for Research and Development in Higher Education, University of California, Berkeley; Center for Social Organization of Schools, Johns Hopkins University, Baltimore, Md.; Learning Research and Development Center, University of Pittsburgh, Pa.; Research and Development Center for Teacher Education, University of Texas, Austin; Stanford Center for Research and Development in Teaching, Stanford University, Cal.; Center for Occupational Education, North Carolina State University, Raleigh; Center for Research and Leadership Development in Vocational Technical Education, Ohio State University, Columbus; and Wisconsin Research and Development Center for Cognitive Learning, Madison.

* * * * * *

THE TEACHER CORPS and the Peace Corps have announced a cooperative plan to train teachers for innovative education programs in the United States and the developing nations. Under the plan, which is expected to include a large proportion of people from minority groups, a young person will spend one year in the Teacher Corps at home and 2 years in a Peace Corps country overseas. The combined service will lead to the masters degree for participants, most of whom will specialize in mathematics and sciences.

Universities and school systems will recruit in their own areas. The Teacher Corps will train the participants for an internship of a year or more in an American school system. The interns will serve in teams under experienced teachers in poverty area schools. This experience and graduate courses will lead to teacher certification at the end of the year.

After certification, participants will be trained for their overseas assignments as Peace Corps volunteers. During this 2-year period, they will assist host countries in planning and implementing innovative education programs. This will complete the requirements for the masters degree in education.

The combined 3-year service is expected to produce a resource pool of innovative specialists and teachers with unique experience for the school systems cooperating in the internship program. The first participants will begin their Teacher Corps service in the summer of 1971 and will be ready for Peace Corps duty in the fall of 1972.

The Peace Corps, created in 1961, has more than 9,000 volunteers and trainees in sixty-one countries. The Teacher Corps, established by Congress in 1965, has 3,200 Corps members serving in thirty-five states, Puerto Rico, and the District of Columbia.

* * * * * *

DURING THE 1970 SUMMER QUARTER Florida State University, through its Council for Instruction, provided support for instructional development. During this period, fourteen faculty members conducted projects related to designing courses, developing teaching skills, using media, and preparing new instructional materials. Two of these projects were conducted by members of the Department of Psychology and one by a member of the Department of Educational Research and Testing.

In the Department of Psychology, Lloyd F. Elfner, conducted a project titled "Research Methodology, Statistical Procedures, and Experimental Psychology." The instructional problem under investigation was the loss of students in Psychology 305 (Statistics for the Behavioral Sciences) which was structured to: (1) provide adequate instruction in research methodology and statistical procedures used in psychology, and (2) provide "visibility" to ways in which skills learned in experimental psychology can be applied in the general field of psychology. It was believed that these objectives resulted in a bored student who would drop out unless he was specifically motivated.

An audio-tutorial approach to teaching the course was undertaken. By integrating visual and audio-channels of information, it was hoped the learning process would proceed at a faster rate with more time available for the student to pursue interests that he found more stimulating. Overall objectives of the course material for the auditory section were spelled out and specific objectives correlated for each tape-slide portion of the course. Materials for the course have been arranged to allow the student to pursue the grade level he wishes to achieve. That is, by passing all the "C" level material the student will earn a "C" in the course. In order to earn a "B," additional material both in terms of tape-slide presentation, laboratory exercises, and outside participation are required. "A" level work in the course is built on further additions to the "B" level materials.

Also in the Department of Psychology, C. H. Madsen, Jr. carried on a project titled "Development of Instructional Video Tapes for Classroom RAID (Rules, Appraisal, Ignore, Disapproval)" presentation. This project dealt with the perennial problem of classroom management. The investigator felt that teachers were often unaware of how their behavior affected children in a classroom. Evidence indicated that attention behaviors given to children may be crucial in the production, maintenance, and/or elimination of disruptive and appropriate behaviors of children.

The project produced a package of training materials which included materials about the following management strategies: (1) making classroom rules repeatedly explicit, (2) ignoring minor behaviors which interfere with learning, (3) using disapproval and isolation when it is necessary to withdraw approval, (4) giving approval, praise, and attention to behavior which facilitates learning, and (5) reinforcing pro-social behavior incompatible with inappropriate social behavior.

In the Department of Educational Research and Testing, John R. Hills carried out a project titled "Revision of EDR 401: Measurement and Evaluation of the Classroom." The instructional problem was to develop an attractive, modern, relevant, and efficient course that would reach a higher percentage of teacher trainees than the current course. Based on interviews with some faculty, department heads, and deans of the College of Education, a list of favorite measurement and evaluation topics was made which included the following: choosing, administering, and interpreting standardized tests, grading, essay tests, and using tests to improve instruction. Based on these topics, a video tape presentation was made, programmed instructional materials located, an item pool prepared, and behavioral objectives developed. Further development and testing of these materials is underway.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

RICHARD R. DEBLASSIE is Head of the Department of Educational Psychology at New Mexico State University. He received the BS and MA from the University of New Mexico, the EdS from New Mexico State University, and the EdD from the University of Arizona. He has had considerable experience as a teacher and counselor at the public school level. Dr. DeBlassie's publications pertain to counseling theory and philosophy as well as to personality variables as they relate to teaching and counseling effectiveness.

RALPH C. DOBBS is Major Academic Adviser in Adult Education at the University of Missouri, College of Education, Department of Higher and Adult Education. He is an experienced teacher at the elementary, secondary, and adult levels and has done considerable research in the analysis of adult learning needs at the community level. He is the author of a textbook entitled *Adult Education in America: An Anthological Approach*, which is being published by Joplin Press. He holds a PhD from Indiana University.

RICHARD A. GORTON is Assistant Professor of Educational Administration and Supervision at The University of Wisconsin—Milwaukee. He received the BA and MA at the University of Iowa and the EdD from Stanford University. He has been a social studies teacher, guidance counselor, assistant principal, and principal. In addition to introducing modular scheduling, team teaching, and several other innovations in administrative reorganization on the secondary school level, he has consulted with school districts and has written articles on innovations in education and student activism.

A. RALPH HAKSTIAN is Assistant Professor of Education at the University of Massachusetts. He received his BA from the University of British Columbia and his MA and PhD degrees from the Laboratory of Educational Research, University of Colorado. Prior to graduate study, he taught high school English and mathematics and gained experience in computer programming. His professional interests include statistics, experimental design, psychometrics, and factor analysis.

ABDUL WAHAB KHAN is a Reader in the Faculty of Engineering, West Pakistan University of Engineering and Technology, Lahore, West Pakistan. He received his BA and BSc from the University of the Panjab and his MSc from Oklahoma State University. He received practical industrial training in Sweden, Germany, and the United Kingdom. He has served as Head of the Department, at Government Polytechnic Institute, Rawalpindi. He has served as member and chairman of a number of educational committees. He has contributed in the "Manual of Standards for Polytechnic Institutes" and is author of more than a dozen of papers on vocational guidance, psychological testing, industrial accidents, technical education, and operational research.

G. BENJAMIN LAWRENCE is Director of the Planning and Management Systems Program and Associate Director of the Western Interstate Commission for Higher Education (WICHE), Boulder, Colorado. He has served as Executive Director, Oregon State Educational Coordinating Council, Portland. From 1956 to 1966, except for a 2-year study leave, he held various executive positions in Cameroon, West Africa for the Baptist Mission Schools. He received a postgraduate certificate and academic diploma in Education from the University of London. He received his MEd and PhD from the University of Oregon. **CLIFTON C. LOVELL** is presently a Research Associate with WICHE, Division of Planning and Management Systems. He previously was on the research staff of the Rand Corporation and also served 6 years as an Air Force officer, primarily involved in systems analysis and operations research. He is a graduate of the Air Force Academy and holds an MS from the University of Colorado and an MA from the University of Southern California. **EDWARD LYELL** is currently a graduate associate with WICHE. He received the BA and MBA from San Francisco State College.

DEANE D. NELSON is Director of the Counseling Center and Associate Professor of Education at Moorhead State College, Moorhead, Minnesota. He received his BS from Gustavus Adolphus College in Minnesota, his MA from the College of Saint Thomas in Minnesota, and his EdD from the University of South Dakota at Vermillion. His professional experience includes public school, college, and university teaching and counseling. He has also been the Director of Counseling at the University of South Dakota. He is a member of APA, APGA, and PDK.

T. N. RAINA is currently Reader in education in the Regional College of Education, Ajmer, India. He was formerly Professor of Education in the Teachers College, Sardarshahr. He received his MA from the University of Panjab and MEd from the University of Delhi. **M. K. RAINA** is Lecturer in the College of Education, University of Indore, India. He was formerly in the Teachers College, Udaipur. He earned his BA, MEd, and PhD from the University of Rajasthan, Jaipur. He has worked as University Grants Commission research fellow in the Regional College of Education, Ajmer.

LOUIS ROSS is an Associate Professor of Mathematics at the University of Akron. He received has BS, BA, and MAEd from the University of Akron and his PhD from Western Reserve University. He developed and taught the first algebra-trigonometry course by closed-circuit TV at the University of Akron. His professional background includes teaching at the undergraduate and graduate levels. Currently he is chairman of the Statistics Section of the American Society for Quality Control. His major areas of interest are mathematics and statistics.

HUGH TAYLOR is an Assistant Professor of Education at the University of Victoria and teaches courses in statistics and evaluation. He received his BA from the University of British Columbia and earned a MEd from Oregon State University and an EdD from Washington State University. His professional background includes teaching mathematics and science at the junior high level, counseling, and pupil personnel work. His primary interests are in educational evaluation and research.

JULIA VANE received her PhD in Psychology from New York University. She is a professor at Hofstra University and Director of the Doctoral Programs in Psychology. She is a consultant to several Long Island, New York, school districts and to the Office of Economic Opportunity in Nassau County. She has published a number of articles on psychological tests and achievement of school children from low socioeconomic backgrounds. **WILLIAM DAVIS** received his PhD in Psychology from Hofstra University. He is Director of Research for the Health and Welfare Council of Nassau County, New York. Prior to this he was Research Director for the Economic Opportunity Commission of Nassau County. He has taught special courses at Hofstra and has served as a Counselor for the educationally disadvantaged students in the New Opportunities Program at Hofstra. His major interests are in the area of research with the disadvantaged.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor of Curriculum and Instruction, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Associate Professor of Special Education, Yeshiva University, New York.

GEORGE S. JURATA, Coordinator of Reading, Jefferson County Schools, Colorado.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

S. JAY SAMUELS, Associate Professor of Educational Psychology, University of Minnesota, Minneapolis.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Editors Emeriti

HARRY J. BAKER, 2241 Q Via Puerta, Laguna Hills, California 92653.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida 32601.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

Diary No. 308
Date 21. 7. 71
library
Bureau Edl. Psy. Research.

51st Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **APRIL 1971** **NUMBER 8**

ARTICLES

Counselor Education: A Profession on the Move—*Thelma M. Urbick; Douglas R. Gross* 339
Flexibility of Cognitive Style and Its Relationship to Academic Achievement in Fourth and Sixth Grades—*Pearline P. Yeatts; Gerald A. Strag* 345
Study Habit Modification and Its Effect on Academic Performance: A Behavioral Approach—*Donald J. Tosi; Richard D. Briggs; Rosemary M. Morley* 347
Relationships Between Personal Variables of Teachers and Their Perception of the Behavior of School Personnel—*Eldon J. Null* 351
The College Environment: Its Meaning to Academically Successful and Unsuccessful Undergraduates—*Dennis Earl Nelson* 355
Assessment of Efforts to Aid Socially Isolated Elementary School Pupils—*Merl E. Bonney* 359
Effect of the System Approach on Changing Preservice Teacher Attitudes Toward Selected Instructional Design Factors—*Thomas G. Kennedy; Robert A. Humphrey* 365
A Study of Children's Drawings of Their Classrooms—*R. S. Rogers; E. N. Wright* 370
An Empirical Comparison of Selected Enrollment Projection Formulas—*William J. Webster* 375

DEPARTMENTS

Among Those Present 384
Book Reviews—*Robert E. Clasen* 35
Field News—*Willard E. North* 38
Forthcoming Articles 38
It Says in The JER 33
New Books from the Publishers 36
Research Briefs—*Harold R. Strang* 34

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (t May-June and July-August issues being combined). The subscription price is $10.00 a year, plus postage for all subscriptions outside the continental United States. Single copies $1.50. Second cl postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, I

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, Preside Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assist to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEW Supervisor of Editorial Services.

It Says in The JER...

COUNSELOR EDUCATION, according to Urbick and Gross, has gone through three distinct periods since its beginning. They discuss these periods—definition and evaluation; standardization; and expansion—and give some suggestions for better counseling during the present era (p. 339).

ACADEMIC PERFORMANCE is a function of the interaction of cognitive flexibility and cognitive fluency and specific task requirements, according to the research reported by Yeatts and Strag (p. 345).

COLLEGE STUDENTS exposed to a treatment which combined psychological conditioning with a well-known study technique demonstrated higher GPA's than a comparable control group in a study conducted by Tosi, Briggs, and Morley (p. 347).

CERTAIN PERSONALITY FACTORS of teachers, along with their attitude toward children, are related to the evaluation of certain dimensions of behavior by other school personnel. Null studied 1,691 teachers in Minnesota schools (p. 351).

DO HONOR ROLL and academic probation students perceive the college environment differently? Nelson found that concepts such as grades, myself as a student, and studying yielded the greatest differences between the groups (p. 355).

"IF YOU WOULD HAVE A FRIEND, be one," has never been more than a half-truth, says Bonney. He studied efforts to aid socially isolated elementary school pupils and found that all group differences were insignificant after assessing seventeen different kinds of socializing experiences (p. 359).

PRESERVICE PROGRAMS must prepare teachers to meet the present as well as future needs of education. Kennedy and Humphrey determined the effect of system approach on changing preservice teachers' attitudes toward selected instructional design factors (p. 365).

CHILDREN'S DRAWINGS of their classrooms showed distinct changes as the children got older. However, there were less dramatic differences between those who attended junior kindergarten and those who did not, according to Rogers and Wright (p. 373).

ENROLLMENT PROJECTIONS for elementary school were more accurate than the projections for secondary enrollment in Webster's study. One of the causes for this is that enrollment in secondary school is further removed from birth than enrollment in elementary school (p. 375).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Editorial/Abstract

The State of the Art:
Counselor Education: A Profession on the Move

At a point in time when counselor education programs are in a state of flux, it seems relevant to pause for a moment of reflection, evaluation, and tentative prediction. Counselor education programs have come a long way from their early beginnings at the turn of the century. The progress has been slow but constant. From its beginnings, counselor education has gone through three distinct periods: (1) definition and evaluation; (2) standardization; (3) expansion. The first two stages exemplify the past and the current status of counselor education programming. The third period, expansion, is a view of the future and the trends that counselor education will follow in its search for relevance.

The counselor must become more involved in the environment of the student; specialists in the pupil personnel team must work more closely together; counselors need more effective preparation for handling current problems of drugs, intergroup relations, and conflicting moral values; career guidance must become developmental; paraprofessionals must be used in support functions, and counselor educators must themselves remain relevant.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

COUNSELOR EDUCATION: A PROFESSION ON THE MOVE

THELMA M. URBICK
Western Michigan University, Kalamazoo

DOUGLAS R. GROSS
Arizona State University, Tempe

IN TERMS OF educational development, the concept of the counselor and, more specifically, counselor education are relatively recent. The first graduate course was offered by Harvard University during the summer session of 1911 (12). Subsequently, other universities began adding one or more courses until, in 1967, 336 colleges and universities provided graduate courses leading to professional degrees in counselor education (21). The growth is phenomenal and the reasons varied. This phenomenon of rapid development is even more pronounced when one realizes that much of this growth had its impetus within the last 2½ decades. This point is illustrated in an article by Kremen in which he stated:

> *The number of counselors in public secondary schools increased two to three times in the period 1939–1945. The number of schools employing counselors tripled during this same period. With the ending of World War II, the growth of programs was further accelerated so that the numbers of counselors have at least doubled again. . . . The May, 1941, issue of* Occupations *listed colleges in the United States, Hawaii, and Puerto Rico which would offer guidance courses in the summer of 1941. Only fifty-one colleges were listed in all. By 1947, this number had increased to several hundred and by 1949, had risen to 980. . . . Thirty-nine institutions offered the Master's Degree in 1946, sixty-four in 1947, and ninety-one in 1949. Sixteen schools offered the Doctorate in guidance in 1946, twenty-seven in 1947, and thirty-nine in 1949 (23:584).*

When compared with today's figures and projected estimates, these totals do not seem too striking, but they do indicate that in a period of a few years, from 1941 to 1949, the number of schools offering guidance courses multiplied nearly twenty times. It is the authors' contention that the vital development of programs of counselor education had its initial impetus during the period of the 40's and early 50's, and that the years which followed have been characterized by attempts at reevaluation, reconceptualization, experimentation, standardization, and expansion. This contention seems to be supported by the apparent lack of literature dealing with counselor education prior to 1940 and the increasing emphasis being placed upon the area in the years that followed. Therefore, certain key articles which are representative of the past 2½ decades of program development have been selected (1, 2, 4, 5, 6, 7, 8, 10, 13, 18, 20, 21, 26, 27, 30, 33, 38, 40, 41, 43, 44). These articles are illustrative of the developmental patterns which have characterized the growth of counselor education programs.

A perusal of these materials suggests certain emerging patterns, and, for the purpose of this discussion, this paper has been divided into three distinct segments: (1) definition and evaluation; (2) standardization; and (3) expansion. Such a division seems to correspond roughly to the period under question, namely the 40's–50's, 60's, and 70's, and provides a meaningful guideline for a discussion of the past, present, and future of counselor education.

PAST

Period of Definition and Evaluation

The definitive period of the 40's and early 50's saw professional organizations and leading counselor educators attempt to define and delimit aspects of counselor duties, qualifications, and training. There was emphasis on methods, techniques, and courses of study (7, 13, 18, 27, 30, 38, 41, 42, 43). Emerging or reemerging patterns seemed to center upon the counselor as an educator, working within the context of an educational setting, as a member of an educational team. Emphasis was placed on the technical aspects of training with emerging emphasis upon the counselor as a technician.

In the decade which followed, evaluation was the key word. It was a period characterized by the testing of the defined or redefined concepts and an introduction of the new. It was a period of experimentation, hypothesizing, and fervent discussion of opinions, concepts, and assumptions. It was a period of evaluation of the old and trying of the new.

Emerging patterns indicated a trend away from the technique oriented program toward one emphasizing attitudinal orientations of counselors. There seemed to be an emerging pattern of professional education, emphasizing psychological and sociological knowledge and skills. Emphasis was placed on counselor competencies, and the organization of programs around these defined competencies emerged. Flexibility and creativity were stressed, and, as the decade progressed, there appeared to be a growing agreement as to the desired characteristics of professional preparation programs.

PRESENT

Period of Standardization

The period of the 60's brought a growing trend toward standardization. Through the efforts of professional organizations, standards were proposed, reviewed, discussed, and adopted (1, 2, 3, 5, 6, 8, 10, 20, 28, 44). These standards provided needed guidelines and direction and seemed to synthesize the thinking and deliberation of both the past and present. They set forth in both specific and general terms concepts regarding the selection, training, and endorsement of counselor candidates. Inherent in such standards was the underlying assumption of adequate preparation for adequate professional participation.

Throughout all of these periods, as noted from the articles, there appeared to be an emerging commonality not only in the recommendations set forth regarding areas of preparation but also in areas of counselor duties, counselor qualifications, and counselor competencies. In fact, certain commonalities are so pronounced that one might be led to assume that they were written by the same individuals at a specific point in time. In reality, however, one realizes that a time span as great as 19 years separates several of these publications.

An analysis of these articles lends itself to conjecture and interpretation. Due to the differing formats, terminology, and inclusions, it is difficult to deal in specifics. Therefore, the interpretations which follow are based on inferred generalizations which the authors found in each of the publications.

One of the significant areas of commonality is that of "counselor role." In those articles which dealt specifically with this topic, it was easy to draw similar comparisons. In articles not dealing specifically with the topic, one had to interpret from other statements and recommendations. In synthesizing the articles, one arrives at the following divisions: assisting the individual, using referral sources, aiding in the development of decision-making processes, presenting relevant information, and using ethical practices.

A second area of similarity centers upon the "counselor as a person." In certain of the publications, this area was referred to as "counselor characteristics," in others "personality factors." In the synthesis, however, the following areas were of a pervasive nature: belief in the individual, commitment to individual human values, alertness to the world, open-mindedness, understanding of self, and professional commitment to the role of the counselor.

A third, and perhaps an area showing the greatest degree of commonality, is that listed under the varying titles of "counselor competencies," "professional studies," and "areas of preparation." This area dealt with the recommendations for content inclusion in the program of studies for counselors-in-training. Within the context of the publications, the following ten areas are easily identifiable:

(1) *Background in the social and biological sciences*
(2) *Practical experience (laboratory work, practicum, internship)*
(3) *Principles and practices of guidance*
(4) *Counseling theory and techniques*
(5) *Organization and administration of a guidance program*
(6) *Knowledge of tests and measurements*
(7) *Knowledge of statistics and research methods*
(8) *Knowledge of individual analysis*
(9) *Knowledge of occupational information*
(10) *Knowledge of group work*

It is interesting to note that, even though as many as 19 years separated the first from the last of these publications, there still appeared to be a great deal of consistency regarding the make-up of an adequate program for the preparation of counselors-in-training. This same consistency is still evident in many of the programs in 1970.

Currently, counselor education is in a state of flux. With increasing demands for relevance in curriculums, emerging emphasis on group work, increasing governmental involvement in educational programs, social movement by minority groups, and changing attitudes within its student population, the horizons for counselor education are somewhat cloudy. The current philosophy seems to be one emphasizing innovation with the credo reading "turn on, tune in, or drop out."

FUTURE

Period of Expansion

The demand for counselors since 1958 has resulted in accelerated preparation programs. The present finds the supply more nearly equaling the demand, and education programs are now focusing on producing more knowledgeable, more effective, and more skillful counselors. Two-year counselor education programs will become the rule rather than the exception (36). Practicum type experiences starting with micro-sessions, tapes, and films will be incorporated from the outset of preparation. A substantial portion of the 2-year preparation will place the neophyte counselor in field placements and internships providing him with essential environmental reality orientation and experience. This could mean an opportunity to explore the possibilities of counseling in different settings for the individual who has not settled on a particular area, it could mean more concentrated depth experience for others (15, 16, 22, 29, 37).

Present trends in counseling indicate that the interdisciplinary involvement will increase and will, therefore, become a greater force in counselor education. If the counselor is to become the effective agent of change, the future will demand that he become more involved in the environment of his counselees, be it the street corner or the home. Psychological, sociological, and anthropological background knowledge will be invaluable in preventing his making initial *faux pas* in these relationships (31, 47).

Because of the press of time and constant demand for deeper knowledge of human development, response, and potential, it will behoove the members of the pupil personnel team—counselor, psychologist, diagnostician, school social worker, nurse, and reading and speech specialists—to work more closely together to evolve answers on mutual student problems. A common core course based on child development and personality theory as well as other shared courses and seminars in psychology, sociology, and anthropology will result in far more effective dialogue in the working environment (9, 11, 24).

If counselor education programs are to effectively prepare future counselors, a real necessity will be a sensitivity to rapid societal changes and the ability of the counselor education staff to make swift curricular changes that take these into account. The present crisis in racial relations, drugs, and conflicting moral values is finding the majority of counselors unprepared to handle the resulting issues (19, 45, 46).

Predictions for the future indicate a population polarization—large percentages of people under 25 and over 60 years of age. Realizing that potential human problem areas are identifiable at an early age, logically counselor education programs will continue to be responsive to the real need for more counseling at the elementary level than presently exists (2, 6, 14, 32, 34, 35). Because the elementary counselor must work very closely with parents and teachers, practical experience with these contacts will likely become an integral part of the supervised experience of the neophyte elementary counselor.

The opposite end of the population polarization prediction—the "over 60 group," may possibly lead to new demands on counselor education programs—that of providing both didactic and experiential learning involvement for counselors wishing to specialize in geriatric counseling. With retirement coming increasingly early in the individual's life, the counselor may have to prepare himself to be an avocational specialist.

Increased use of the computer's superior memory for vocational and education information will be an important factor in the decade of the '70's. Inasmuch as the counselor will not simply be an information dispenser, his time and talents can be more efficaciously used as a facilitator in the decision-making process requisite to career determination (25).

Because of the increasing need of an increasingly complex society to utilize to the utmost the potential of each individual, career guidance will, of necessity, become developmental starting with orientation and exploration in the elementary and junior high school years. At present, the majority of counselor education programs do not provide counselors with the background for adequate leadership in career development. Too often looked upon as the "second class citizen" in counselor education, the classes in vocational and educational information are taught and endured rather than challenged and stimulated. This, in the future, will hopefully be envisioned as the important area it is and as an area that merges with the entire counselor education curriculum through practicum type experiences, i.e., counseling sessions with individuals puzzling over vocational direction for their lives.

Rogers (39) has indicated that the future necessity of greater numbers of people living within a circumscribed amount of space is being manifested in the increased emphasis on and interest in encounter groups. Learning to live in a group-like situation, according to Rogers, will be important for real and fulfilling relationships. An understanding of group dynamics and experiences in being both a group member and a group leader (which will often include a practicum in group counseling) will become a part of all on-going counselor preparation programs (17).

Increasing use of the paraprofessional (individuals without professional preparation but trained to perform certain support functions) will relieve counselors of duties not requiring a high level of professional skill. It will also provide work opportunity for people who can serve as liaison between the counselor and a unique population, e.g., minority groups with whom he may have difficulty establishing an initial relationship. Counselor educators will undoubtedly be responsible for in-service type orientation programs to prepare the paraprofessional for his role.

Utilization of paraprofessionals will have still another influential impact. In freeing counselors from non-professional level involvement, it will allow time to be spent in implementing one of the real goals of counseling—helping people to gain more realistic self-concepts. With future predictions indicating a growth in professional and technical workers, continued advances in automation and cybernation, and concomitantly, an increasing number of jobs that cause the worker to become a pusher of buttons, a turner of wheels, a setter of gauges, it follows that it will be increasingly difficult for work to be the focus of life and the prop of the self-concept. Counselor education, therefore, will need to provide a far better background of knowledge and experience on the development and implementation of the self-concept. All facets of preparation programs will require orientation in this direction with consequent emphasis on the areas of experiential involvement and expertise in effective supervision.

Inasmuch as the counseling profession cannot logically require more of its counselors than it does its counselor educators, these individuals, too, will be faced with greater expectations, didactically, experientially, and personally. More reading, more seminars and workshops, can assist the educator with "textbook knowledge." However, the counselor educator who does not also keep himself involved in counseling, and through observation and experience, aware of the real problems facing counselors in various settings, may find himself faced with justifiably derisive students. This is true also in the personal area—the demand increases to *be* what you *say*. The counselor educator faces challenging requirements; his learning, his experience, and his development as a real human being will never cease if he is to be a dynamic part of an on-going, effective, truly professional education program.

REFERENCES

1. American Personnel and Guidance Association, "The Counselor: Professional Preparation and Role: A Statement of Policy," Author, Washington, 1964.
2. American Personnel and Guidance Association, "The Counselor: Professional Preparation and Role: A Statement of Policy," Author, Washington, 1967.
3. American Personnel and Guidance Association, "A Report of the Committee of Professional Preparation and Standards," Author, Washington, 1963.
4. American Personnel and Guidance Association; Association for Counselor Education and Supervision, "Standards for the Preparation of Elementary School Counselors (1st draft)," Washington, February 1962.
5. American Personnel and Guidance Association, "Standards for Preparation of School Counselors," Author, Washington, 1961.
6. American Personnel and Guidance Association, Association for Counselor Education and Supervision, "Standards for the Preparation of Secondary School Counselors—1967 Edition," Author, Washington, 1967.
7. American Psychological Association, Division of Counseling and Guidance, Committee on Counselor Training, "The Practicum Training of Counseling Psychologists," *American Psychologist*, 7:182–188, 1952.
8. Arbuckle, D. S., *Counseling Philosophy Theory and Practice*, Allyn and Bacon, Boston, Massachusetts, 1965.
9. Arbuckle, D. S., "Counselor, Social Worker, Psychologist: Let's Ecumenicalize," *Personnel and Guidance Journal*, 45:532–538, 1967.
10. Berdie, R., "Who is Responsible for Preparing College Student Personnel Workers?" paper presented at the American Personnel and Guidance Association Convention, Dallas, Texas, March 1967.
11. Bertness, H. J., "School Counseling and the Pupil Personnel Services Movement," in Stiller, A. (ed.), *School Counseling, 1967, A View from Within*, Yearbook of the American School Counselor Association.
12. Brewer, J. M., *History of Vocational Guidance*, Harper, New York, 1942.
13. Burnett, C. W., "Selection and Training of School and College Personnel Workers," *Review of Educational Research*, 24:121–133, 1954.
14. Cottingham, H. F., "Conceptual Elements of Elementary School Guidance," (initial draft), unpublished report, Florida State University, Tallahassee, 1967.
15. Dunlop, R. S., "Pre-practicum Counselor Education: Use of Simulation Programs," *Counselor Education and Supervision*, 7:145–146, 1968.
16. Farwell, G. F., "The Introductory Course: A Synthesis of Theory and Practice," *Counselor Education and Supervision*, 7:98–101, 1968.
17. Foreman, M. C., "T Groups: Their Implications for Counselor Supervision and Preparation," *Counselor Education and Supervision*, 7:48–53, 1967.
18. Froehlich, C. P., "Content of the Manual on Counselor Preparation," *Occupations*, 27:541–545, 1949.
19. Herr, E. L., "The Impact of Recent Student Protests and Student Movements: Implications for Counselors and Counselor Educators," *Counselor Education and Supervision*, 6:236–247, 1967.
20. Hill, G. E.; Green, D. E., "The Selection, Preparation, and Professionalization of Guidance and Personnel Workers," *Review of Educational Research*, 30:115–130, 1960.
21. Houghton, M. W.; Trexler, L. M., *Inventory of Counselor Education Programs 1965–66*, U. S. Government Printing Offices, Washington, D. C., 1967.
22. Joslin, L. C., Jr., "The Introductory Course: A More Meaningful Emphasis," *Counselor Education and Supervision*, 7:93–97, 1968.
23. Kremen, B. G., "Counselor Certification in the United States," *Occupations*, 29:584–586, 1951.

24. Liddle, G. P., "Interprofessional Relationships Among the Pupil Personnel Services," paper read at Pupil Personnel Council, Kent State University, Kent, Ohio, April 1962.
25. Loughary, J. W., "Will Counseling be Eliminated in the Cybernation Revolution?" *Preparing School Counselors in Educational Guidance*, College Entrance Examination Board, New York, 1967.
26. MacMinn, P.; Ross, R. G., *Status of Preparation Programs for Guidance and Student Personnel Workers*, Bulletin No. 7, United States Department of Health, Education and Welfare, 1959.
27. Mathewson, R. H., "Graduate Training for School Counselors," *Occupations*, 31:335–339, 1952.
28. Mathewson, R. H., "Some Observations on Research in Counselor Preparation," paper read at Conference on Research in Guidance and Counseling, Georgia University, Athens, January 1967.
29. Meek, C. R.; Parker, A. W., "Introductory Counseling Course: Use of Practicum Students," *Counselor Education and Supervision*, 5:154–158, 1966.
30. Miller, R. W., "Counselor-training Programs in Colleges and Universities," *Personnel and Guidance Journal*, 32:132–134, 1953.
31. Miller, R. W., "The Interdisciplinary Nature of Counselor Education," paper read at the 31st Annual Conference on Guidance and Personnel Work, Northwestern University, Evanston, Illinois, July 1967.
32. Muro, J. J.; Oelka, N. C., "Guidance Needs in the Elementary School, Due to the Preparation of Counselors," *Counselor Education and Supervision*, 7:7–12, 1967.
33. National Vocational Guidance Association, *Counselor Preparation*, 1949.
34. Nelson, R. C., "The Preparation of Elementary School Counselors: A Model," *Counselor Education and Supervision*, 6:197–200, 1967.
35. Nelson, R. C., "The Millenium: Preparation Programs for Elementary School Counselors of the Future," paper read at the American Personnel and Guidance Association Convention, Dallas, Texas, March 1967.
36. Noble, F. C., "The Two-year Graduate Program in Counselor Education: A Re-examination," *Counselor Education and Supervision*, 4:160–162, 1965.
37. Perrone, P. A.; Sanborn, M. P., "Early Observations: An Apprenticeships Approach to Counselor Education," *Counselor Education and Supervision*, 6:63–68, 1966.
38. Pierson, G. A., "Utilizing Internships in the Preparation of Counselors," *Occupations*, 29:414–419, 1951.
39. Rogers, C. R., *Freedom to Learn*, Merrill Publishing Company, Columbus, Ohio, 1969.
40. Tooker, E. D., "Counselor Role: Counselor Training," *Personnel and Guidance Journal*, 36:263–267, 1957.
41. United States Office of Education, Federal Security Agency, "Duties, Standards and Qualifications of Counselors," Division of Vocational Education, Miscellaneous 3314–1, Superintendent of Documents, Government Printing Office, Washington, D. C., 1949.
42. Wrenn, C. G., "Training of Vocational Guidance Workers," *Occupations*, 29:414–419, 1951.
43. Wrenn, C. G., "Selection and Education of Students' Personnel Workers," *Personnel and Guidance Journal* 31:9–14, 1952.
44. Wrenn, C. G., *The Counselor in a Changing World*, report of the Commission on Guidance in American Schools to the American Personnel and Guidance Association, The Association, Washington, D. C., 1962.
45. Wrenn, C. G., "A Second Look," in Loughary, J. W. (ed.), *Counseling a Growing Profession*, American Personnel and Guidance Association, Washington, D. C., 1965, pp. 55–56.
46. Wolfbein, S. L., "The Impact of Social and Economic Change," *Counselor Education and Supervision*, 6:230–235, 1967.
47. Zaccaria, J. S., "Guidance Implications of Concepts from the Field of Culture and Personality," *Personnel and Guidance Journal*, 45:907–910, 1967.

Research Briefs

SOME PRELIMINARY FINDINGS ON TOKEN SPENDING PATTERNS IN DISADVANTAGED GRADE SCHOOL STUDENTS[1]

WITH THE widespread use of token economies in educational settings, increased emphasis has been placed upon extremely elaborate reward systems allowing students to redeem tokens for a wide variety of items. Such a trend is typified in the Multiple Reinforcer Rule by Ayllon and Azrin (1). The efficiency of such elaborate systems, however, is questionable since little empirical data concerning children's spending patterns within multiple reward systems have been reported.

Some preliminary findings relating to children's reward choices have been derived from an automated remedial reading project conducted at Jefferson Elementary School in Charlottesville, Virginia, during the 1969–1970 school year (2). Twenty sixth-grade students from disadvantaged homes (13 boys, 7 of whom were Negro and 7 girls, 3 of whom were Negro) who had been diagnosed poor readers, received tutoring in reading. During the program, points were awarded for accurate reading. For every 25 points earned a student received a token valued at $.05. Tokens could be saved and later spent for high value rewards from local stores (items of a value > $.50) or redeemed daily for low value rewards (e.g., candy, cookies, or potato chips).

Results and Discussion

During the tutoring students averaged an equivalent of $.07 per student per day in earnings. Female average earnings were only 2.5 percent higher than male average earnings. Of the 3,298 tokens earned, 71 percent were spent for short term rewards. Sixteen of the twenty students spent more for low value than for high value rewards ($\chi^2 = 7.2$, $p < .01$). Also, children chose significantly more candy than the sum of all other low value rewards $\chi^2 = 61.9$, $p < .01$).

With respect to high value spending, analyses across sex revealed that while only five of thirteen boys invested in any high value rewards, a significantly greater ratio of girls, six of seven, did (Fischer exact probability = .05). The average high value girl's spending of thirty-six tokens for the year was 80 percent higher than the average high value boy spending. Even though the ten Negro students earned 10 percent more tokens than did the ten white students, white students spent 32 percent more tokens on high value rewards than did the Negro students ($\chi^2 = 18.92$, $p < .01$).

While on a superficial level, the students' selection of numerous types of rewards (over 40 different high and low value rewards) supports the Multiple Reinforcer Rule, analyses of selection frequencies revealed that a relatively small number of low value rewards dominated spending. Not only did three types of candy account for over 70 percent of low value token spending, but five specific low value rewards accounted for over 66 percent of

(Continued on page 364)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Flexibility of Cognitive Style and Its Relationship to Academic Achievement in Fourth and Sixth Grades

PEARLINE P. YEATTS and GERALD A. STRAG, University of Georgia

ABSTRACT

This research is specifically concerned with whether or not students who tend not to shift cognitive styles are capable of achieving equally as high on verbal and quantitative tasks as students who shift cognitive styles. The Kagan Cognitive Style test was used to identify flexibility and rigidity of cognitive styles. The California Achievement Test was used as our index of verbal and math achievement. Sixty-two fourth grade students and fifty-nine sixth grade students from a north Georgia elementary school constituted the example. Of the thirty-one subjects, twenty-five identified as inflexible and were performing below grade level, only six were performing at grade level, and none were performing above grade level. Data suggest it is inappropriate to consider one cognitive style superior to another. It is evident in the light of the findings of this study to consider academic performance a function of the interaction of cognitive flexibility and cognitive fluency, and the specific task requirements.

LEARNING STYLES are personal ways in which individuals process information in the course of learning concepts. During the past 10 years a number of researchers have inferred that cognitive styles are reflective of individual predispositions to perceive and respond to stimuli in particular ways (1, 2, 3).

Kagan, Moss, and Sigel identify two basic styles of processing information which they refer to as: "analytic and nonanalytic." Individuals who are analytic are predisposed to arranging aggregates of symbols, signs, or figures into groups. This grouping process is carried on by the individual who identifies particular characteristics which appear to be common to the group. Individuals who appear to be nonanalytic are less attentive to specific characteristics and are inclined to be diffuse and global in their observations. In all, most Ss are capable of making use of both types of cognitive responses when and if they are pressed to do so (3).

The focus of this study is upon the relationship between an individual's ability to behave flexibly (to shift his cognitive style) and his academic achievement as measured by a standardized achievement test. The research is specifically concerned with testing the hypothesis of whether individuals who tend not to change their cognitive style are equally capable of scoring above their grade level, on verbal and quantitative tasks, as those individuals who demonstrate flexible cognitive styles.

Procedure

Subjects. The Ss comprised the entire fourth (N = 62) and sixth (N = 59) grade populations of a north Georgia elementary school.

Instruments. The two instruments used in the investigation were the Kagan Cognitive Style Test (KCST) (2) and the California Achievement Test. The criteria of cognitive style is defined in terms of the student's performance as defined by the KCST.

The students were individually tested on all nineteen test items. Each item in the KCST allowed the S 45 seconds to arrange the materials in as many ways as he was able. The instructions were standardized as follows: ". . . put the task together in as many ways as possible."

The student's first answer on each of the nineteen items identified his cognitive style preference. Successive responses to the same item were analyzed as to sameness or difference of cognitive style. Each student was then assigned a flexibility

Table 1.—Relationship Between Flexibility of Cognitive Style and Academic Performance in Fourth Grade

Group	Above Grade	At Grade	Below Grade	Total
As	13	10	0	23
ANs	0	1	6	7
NAs	12	11	1	24
NANs	0	2	6	8
Total	25	24	13	62*

*chi-square = 54.2; p = .01.

As = Analytic Ss who shift
ANs = Analytic Ss not shifting
NAs = Non-analytic Ss who shift
NANs = Non-analytic Ss not shifting

Table 2.—Relationship Between Flexibility of Cognitive Style and Academic Performance in Sixth Grade

Group	Above Grade	At Grade	Below Grade	Total
As	14	7	1	22
ANs	0	1	8	9
NAs	9	10	2	21
NANs	0	2	5	7
Total	23	20	16	59*

*chi-square = 47.0; p = .01.
Group abbreviations are same as for Table 1.

score and a fluency score. The flexibility score consisted of the number of times the student changed his cognitive style. The student's fluency score was the sum of all responses made to each of the nineteen test items.

With regard to the students' achievement scores in mathematics and verbal ability on the California Achievement Test, the following criteria were arbitrarily set. Students who scored no more than 6 months above or no less than 6 months below grade level were classified as "at grade level." Those students scoring less than 6 months below grade level were classified as "below grade level"; wheras those students scoring more than 6 months above grade level were classified as "above grade level."

Results and Discussion

Students whose cognitive style vacillated between analytic and nonanalytic, as exemplified by their responses, were classified as "flexible," while those students not shifting cognitive styles were classified as "inflexible." The students were ranked in order of performance on the task; chi-square was then used to determine if the pattern of placement was a function of chance. The results were significant beyond the .01 level, for both the fourth and sixth grade students (see Tables 1 and 2).

Inspection of the data indicate that students who changed cognitive style performed at a higher level of achievement on the California Achievement Test than the students who did not change their cognitive style. This was the case regardless of the cognitive style preference or the grade level of the student.

Fifteen fourth graders and sixteen sixth graders did not change their cognitive style on these tasks (remained inflexible). Of the thirty-one Ss, twenty-five identified as inflexible were performing below grade level, while only six were at grade level, and none were above grade level.

A multiple correlation between cognitive fluency, cognitive flexibility, and math and verbal performances on each group was calculated. The resulting multiple and partial correlations are given in Table 3.

Table 3.—Multiple and Partial Correlations Among Cognitive Fluency, Flexibility, and Math and Verbal Scores

Group	N	Multiple r Math	Partial Fluency	Partial Flexibility	Multiple r Verbal	Partial Fluency	Partial Flexibility
Fourth Grade	62	.82*	.29	.70*	.80*	.69*	.29
Sixth Grade	59	.89*	.24	.72*	.87*	.69*	.27

*p = .01.

The data in Table 3 suggest that achievement in mathematics and verbal tasks were functionally interacting with cognitive fluency and cognitive flexibility. This was the case with both grade levels tested. In other words, the data clearly suggest that students who are more *fluent* do better in verbal tasks, while *flexible* students do better in math.

Summary

The data suggest that it is not appropriate to consider one cognitive style preference superior to another. It becomes evident in the light of the findings of this study to consider academic performance as being a function of the interaction of cognitive stlye flexibility and fluency, and the specific task requirements. These results further support the importance of behavioral research from the interactionist approach.

REFERENCES

1. Kagan, J.; Moss, H. A.; Sigel, S. E., "Psychological Significance of Styles of Conceptualization," Monographs of Social Research, *Child Development*, 28: (serial no. 86, whole no. 2) 73–112, 1963.
2. Kagan, J.; Rosman, D.; Day, D.; Albert, J.; Phillips, W., "Information Processing in the Child, Significance of Analytic and Reflective Attitudes," *Psychological Monographs*, 78:(no. 1, whole no. 578), 1964.
3. Bruner, J. S.; Goodnow, J. J.; Austin, G. A., *A Study of Thinking*, John Wiley and Sons, New York, 1956.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Study Habit Modification and Its Effect on Academic Performance: A Behavioral Approach

RICHARD D. BRIGGS
Bowie State College, Bowie, Md.

DONALD J. TOSI
Ohio State University

ROSEMARY M. MORLEY
Kellogg Community College

ABSTRACT

The purpose of this research was to determine the effects of a treatment procedure which combined psychological conditioning with a well-known study technique (RSQ3R) upon the academic performance of "high-risk" students. It was hypothesized that Ss exposed to this procedure would demonstrate higher GPA's than a comparable control group. The results were in the predicted direction.

STUDENT PERSONNEL people in higher education have long been concerned with the problem of what to do with the "probationary" or "high risk" student. Some may argue that these students should not be admitted to college in the first place, or if admitted, and there is no evidence of successful performance, they should be dismissed. It may also be argued that "high risk" students lack certain skills which are usually associated with good academic performance, such as effective study methods and self-control over the study situation. If these conditions could be remedied, a substantial increase in the academic success rate of these students would probably be noted.

Research would suggest that two fundamental operations should be included in programs that are designed to assist "probationary" or "high risk" students (1, 2, 3, 4). *First,* the student should be taught an efficient study technique for assimilating information. *Second,* he should learn sufficient self-control over the study situation to insure the successful implementation of the study technique over an extended period of time. In most cases, failure to engage in study is produced by a combination of the aversive nature of the study situation and competition with various social reinforcers, such as having a good time with friends.

In the present study, a procedure was developed which combined (a) simple psychological conditioning procedures, based on the Premack principle (7), to reduce the aversive properties of the study situation, and (b) The Robinson SQ3R, study technique (10), to insure a more fruitful utilization of the study occasion. It was hypothesized that "high risk" students exposed to this procedure would later demonstrate significantly higher grade point averages (GPA) than a comparable control group.

Sample

Subjects consisted of freshman girls ($N = 20$) selected from a calculated risk population in a small 4-year liberal arts college for women. All Ss were of probationary status prior to their first semester of college.[1] They were divided into two groups, A and B. Group A ($N = 10$) contained Ss who had been previously exposed to a special program during a summer orientation period.[2] These two groups were then randomly assigned into two experimental groups, A_1 and B_1, and two control groups, A_2 and B_2.

Method

Ss were informed by E's that their free hours during the day could be utilized for study so their evenings and weekends would be relatively free of study activities. Before the RSQ3R study technique was introduced, it was necessary to place the initiation of studying under stimulus control. The following procedures were initiated. For each academic course Ss were taking, 1 hour was set aside each day, where they could go to a private room and study. Ss were asked to choose a study room which would be relatively comfortable and free of the usual disturbances. They were also instructed to use the same room each day and upon entering this room, they were to study only materials relevant to a particular subject area. During the first day of the experiment, Ss were to study until they felt a desire to stop or experienced any displeasure at all. At this point, no matter how many minutes passed, they had to leave. Following this initial experience, Ss could only leave the room after certain prerequisites were met. On the second day, the prerequisite was to read one page or complete one problem correctly, according to the subject being studied.

On the third and fourth days of the experiment the prerequisite for leaving the room, after a decision to stop, was two pages or two problems. From the fifth through ninth days, the prerequisite was increased to three pages or three problems. After the ninth day, the prerequisite for leaving was raised to a final level of four pages or four problems. Approximately one week after this final increase, almost all Ss were spending the required hour per class in the study room. (See Table 1 for the $\overline{X}$ time spent in the room during this acclimation period.)

Premack Principle. The procedure of acclimating Ss to the study situation was based on the Premack principle of reinforcement (5, 6). The Premack principle suggests that you reinforce a low probability behavior (such as studying) with a high probability behavior (such as having a good time with friends). In this study, the Premack principle would imply that the act of studying received reinforcement in the form of satisfaction over the completion of an assignment as well as earning the *right* to leave the study situation to engage in other activities. The aversiveness of the study situation was further minimized by giving each S the *option to leave* after completing a specific task. By requiring the Ss to remain only a short time in the study room, conditions were provided which would minimize the competition of social reinforcers and the aversive nature of study. Further aversive features were minimized when the option to abandon study was made an approved procedure rather than a display of student irresponsibility.

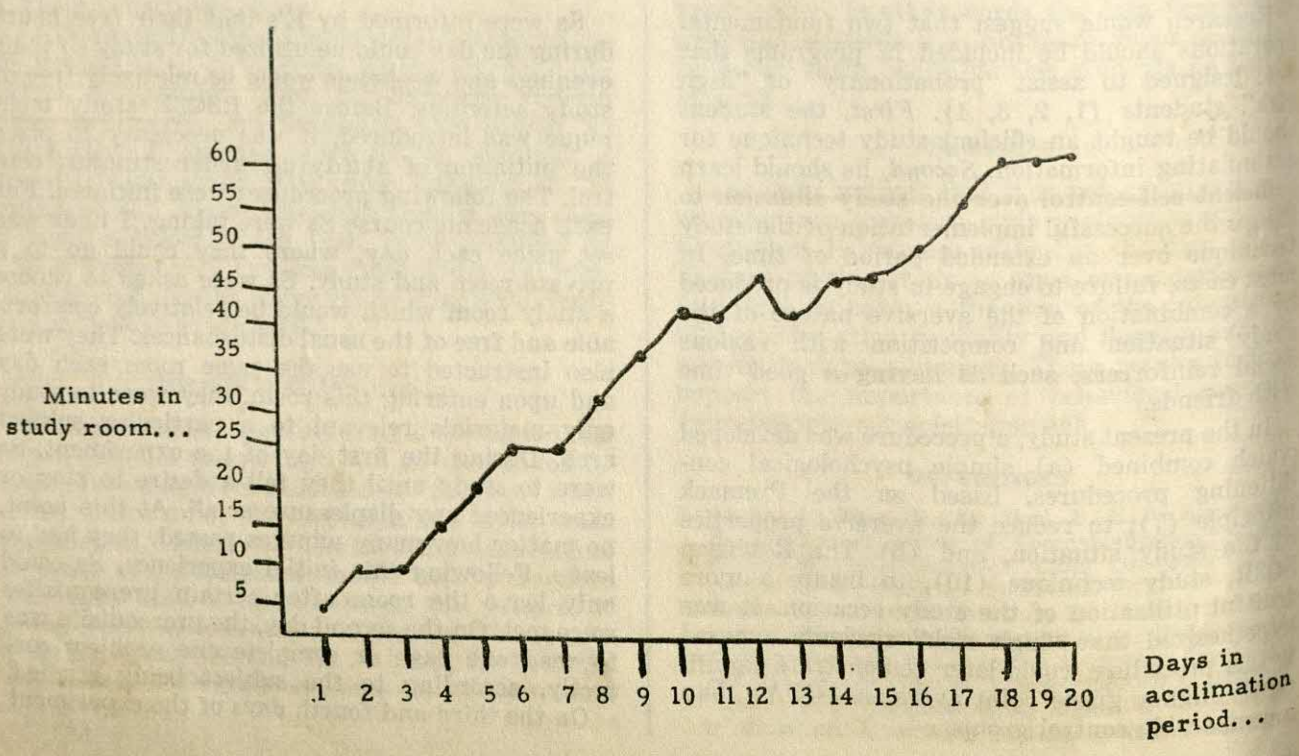
The conditioning principle of successive approximation was also employed by requiring the student to master the program of studying in small increments. Technically, the student was put on a fixed ratio schedule of reinforcement. It is well established that such a schedule produces high rates of responding and increased resistance to extinction (8, 9).

In summary, studying was reduced to small ratios and was dependent in part on escape from the aversive character of the traditional conditions of study. The ratio had been increased gradually to promote work at high ratios by slow increments. Finally, these stimuli were controlled so that they served only as occasions for study—study which terminated in reinforcement.

Introduction of RSQ3R. Once the initiation of study was brought under control, and Ss were spending an hour each day for each class they had, then the Robinson SQ3R method was introduced. The title for this higher-level study skill is abbreviated in the current fashion to make it easier to remember and to make reference to it more simple. The symbols stand for the steps which the student follows in using the method. A brief description of each step is given below:

Survey	1. Glance over the headings in the chapter to see the big points which will be developed.

Table 1.—Experimental Group's Mean Average of Time Spent in the Study Room per Class per Day for the 20-day Acclimation Period.



Question 2. Turn the first heading into a question, etc.

Read 3. Read to answer that question.

Recite 4. Having read the first section look away and from memory write as much about the material as you can remember (most important aspect).

Review 5. Fill in points you couldn't remember by going back to the material.

This method was introduced step by step; only after mastering one aspect of the technique were the Ss allowed to go on. The majority of the Ss were using the complete method within 2 weeks of the acquisition period.

In summary, over a 3-week period, the Ss were being conditioned to spend a sufficient amount of time per class by using simple psychological behavior principles. Then, a study technique (RSQ3R) was introduced to insure that Ss were utilizing that time more efficiently.

Statistical Analysis of Data

A two-way fixed effects factorial analysis of variance was used in analyzing these data. Cell entrees consisted of Ss' obtained GPA's (see Table 2).

Table 2.—Obtained GPA's by Various Groups

	Summer		Non-summer	
Experimental	1. 2.07 2. 1.93 3. 2.46 4. 2.07 5. 2.30	Mean GPA.= 2.17 A₁	1. 1.78 2. 3.14 3. 2..75 4. 2.00 5. 2.00	Mean GPA.= 2.33 B₁
Control	1. 1.25 2. 0.83 3. 1.63 4. 2.20 ·5. 2.16	Mean GPA. = 1.61 A₂	1. 1.73 2. 2.50 3. 2.06 4. 1.78 5. 2.20	Mean GPA. = 2.05 B₂

Complete Experimental Mean GPA = 2.25
Complete Control Mean GPA = 1.83
Complete Summer Mean GPA = 1.89
Complete Non-summer Mean GPA = 2.19

Findings and Discussion

As revealed in Tables 2 and 3, the experimental group differed significantly ($p < 05$) from the control group with respect to their obtained $\overline{X}$ GPA's. $\overline{X}$ GPA's for experimental Ss were significantly higher than $\overline{X}$ GPA's for control Ss. These data certainly suggest that the observed difference in the criterion could be reasonably attributed to the experimental treatment. An insignificant F ratio ($p < 05$) with respect to $\overline{X}$ GPA's was observed between Ss comprising the summer and non-summer groups. Thus, when the experimental treatment was ignored, Ss who were previously involved in a special summer program did not differ significantly with respect to their $\overline{X}$ GPA's from Ss who did not have the experience. No significant interaction effect was observed (see Table 3).

The results of this research indicate the feasibility and applicability of combining psychological conditioning techniques with a study technique in terms of its effect upon the academic performance of "high risk" college students. "High risk" students exposed to the treatment procedures described in this study demonstrated significantly higher GPA's than a comparable control group. Thus, the technique shows much promise in ameliorating the faulty study habits usually associated with "high risk" or "probationary" students. It should also be noted that such a treatment procedure is one of relatively short duration, approximately 5 weeks.

While the investigators would like to assert that the observed differences in the criterion were due exclusively to the experimental treatment, such a conclusion at this time would be premature and unwarranted. Before this approach can be recommended for use on a large scale basis, it will be necessary to conduct additional research on this topic with some modification in the present research design. In improving the research design, a third group receiving only programmed instruction could be used. It might also be advisable to include one group who would be conditioned to spend time in the study room and another group

Table 3.—Table of two-way Fixed Effects Analysis of Variance For $\overline{X}$ GPA's.

Source	SS	DF	MS		F
ROWS (Summer-non-summer)	.51	1	.51	16	3.18
COLUMNS (Experimental-Control)	.91	1	.91	16	5.68*
INTERACTION (Sum. Exp. - N. Sum. Exp.) (Sum. Con. - N. Sum. Con.)	.44	1	.44	16	2.75
ERROR (within the cells)	2.93	16	—		—
TOTALS	---	19	—		—

* $p < .05$

that would receive training in the RSQ3R technique only. Comparisons then could be made between these various groups which would permit the investigator to determine whether the criterion is being affected by the combined effect of the conditioning procedures and the RSQ3R technique, or by any of these alone. By employing a group which receives only programmed instruction, one might be able to control for the possible effects of Ss just being involved in a study.

As counselors and educators, we have a responsibility to search for pragmatic answers to the questions posed by our clients and students. The treatment procedures described in this study offer some innovative practices which show promise, at least in terms of a starting point, to deal squarely with the problem of poor academic performance and its residual counterparts.

FOOTNOTES

1. The probation or calculated risk population is defined as having a deficiency in one or more of the following areas required for admissions. Either they scored low on the ACT college entrance exam, had low grades in high school, or were low in regard to rank in class; recommendations and personal interviews were also taken into consideration.
2. Subjects attended a summer remedial program where they were confronted with a learning theory and attended condensed classes in basic skills (English, Math, etc.).

REFERENCES

1. English, H. B.; Welborn, E. L., "Studies in Substance Memorization," *Journal of General Psychology*, 11: 233–59, 1934.

2. Fox, Ljungberg, "Effecting the Use of Efficient Study Habits," *Journal of Mathematics*, 1:75–86, 1962.

3. Good, C. V., "What Little is Learned Seems to be Forgotten so Rapidly," *Journal of Educational Materials*, 5:325–29, 1926.

4. Homme, L., "A Demonstration of the Use of Self Instructional and Other Leading Techniques for Remedial Intructions of Low Achieving Adolescents in Reading and Mathematics," U. S. Office of Education, Contract No. OE–4–16–033, 1964.

5. Homme, L.; C'de Baca, P.; Devine, J. V.; Steinhorst, R.; Rickert, E. J., "Use of the Premack Principle in Controlling the Behavior of Nursery School Children," *Journal of the Experimental Analysis of Behavior* 6:544, 1963.

6. Jex, Frank B.; Reed, Merril, "Intellectual and Personality Characteristics of University of Utah Students," *Journal of Educational Research*, 53:225–260, 1959.

7. Kimble, "Conditioning and Learning," Appleton-Century-Crofts, Inc., New York, 1961.

(ed.), *Nebraska Symposium on Motivation*, University of Nebraska Press, Lincoln, 1965.

9. Reynolds, G. S., *A Primer Operant Conditioning*, Scott, Foresman, and Co., Glenview, Illinois, 1968.

10. Robinson, F. P., *Effective Study*, Harper and Brothers, New York, 1946, p. 29.

BOOK REVIEWS

ROBERT E. CLASEN
Book Review Editor

Modern Movements in Educational Philosophy

Morris, Van Cleve. Houghton Mifflin Company, Boston, 1969, 381 plus xii pp.

THIS IS a compilation of parts of writings of some twenty-three modern thinkers about education, and one article from *Time* magazine. The compiler's preface states that the volume is the result of intrapersonal dialectic. We have to assume that he expects to use the materials to help students develop some philosophic views of their own other than those dictated by ordinary everyday experience. Morris has grouped the writings in four divisions: Uses of Philosophy; Scientific Philosophy in Educational Thought; Philosophical Anthropology in Educational Thought; Analytic Philosophy in Educational Thought.

The philosophical writers he selects from are Isaiah Berlin, Anatol Rapaport, John Dewey (a liberal quantity from him), John A. Kouwenhoven, John L. Childs, Donald O. Herb, Alexander Calandra, Ernest Cassier, David Bidney, Alexander Edel, Theodore Brameld, Abraham Kaplan, A. J. Ayer, John Wilson, Israel Scheffler, B. P. Komisar, and J. E. McClellan, Fernando Molina, Robert G. Olson, Huston Smith, William Barret, Martin Buber, and Van Cleve Morris. This gives a very wide distribution of views about education.

Some of the excellent features of this material are good editing and printing; articles selected deal with critical problems (on thinking, values, man and his culture, the topology of teaching concept, progressive education, and many others). The mere fact that these materials are compiled in a single volume makes it a great convenience. To each writer's contributions he adds a few helpful supplementary references.

The last section, on Existentialism, may even help some of us poor souls who are befuddled by that recent jargon. The reviewer is of the opinion that a considerable number of students absorb some one leader's views, e.g., Dewey's, and blissfully go on assuming that there are no other relevant ideas. The management of a course using such a compilation might be a cure for this rather limited view of education.

The how of the how to teach such a course bothers the reviewer. No matter what material is used, a course in philosophy of education is a difficult course to teach and get any useful results. The reviewer once sat in on a course by a distinguished philosopher, but he judged then and now that not more than five or six of the students had any "respectable" understanding of what was being done. Many of the man's talks "sailed over our heads" and we later corralled three graduate students to give us some help. The reviewer suggests that the use of this volume without much planning, readjustment, and working *with* students can produce a gigantic spread of befuddlement and discourage numerous human beings about education, the noblest of all professions.

Another caution the reviewer will add: No student should accept the classification Morris gives these writers as correct, complete, and final. Dewey is not all scientific; Brameld is not all anthropological, and Green has not said the last word on the topology of teaching.

In the hands of a mature and careful teacher who knows these views and who can make frequent and sudden readjustments for different students with different ideas, the experiences should be exciting, motivating, and satisfying.

Dr. A. R. Mead, State Director
National Retired Teachers Association
Gainesville, Florida 32601

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Relationships Between Personal Variables of Teachers and Their Perception of the Behavior of School Personnel

ELDON J. NULL, Purdue University

ABSTRACT

The purpose was to determine if relationships exist between personal variables of teachers, as measured by the Minnesota Teacher Attitude Inventory (MTAI) and the Sixteen Personality Factor Questionnaire (16 PF), and their evaluation of the behavior of their principal and fellow teachers in responses to items of the Organizational Climate Description Questionnaire (OCDQ). The sample consisted of 1,691 teachers in eighty-one Minneapolis, St. Paul, and suburban elementary schools. Data were analyzed through the use of coefficients of correlation and multiple regression equations. The examination of 138 null hypotheses indicated that certain personality factors of teachers, along with their attitude toward children, are related to the evaluation of certain dimensions of behavior as indicated by scores on the eight subtests of the OCDQ.

FOR MANY years superordinates have exercised the right and responsibility to evaluate the behavior of subordinates in the school organization. Superintendents, for example, have judged the work of their building principals, and the latter, in turn, have evaluated the performances of members of their teaching staffs. These types of evaluations probably will become more numerous and comprehensive in future years because of the increasing demand by the general public for "accountability" concerning the manner in which tax monies are being used for education. In addition, subordinates now are judging the behavior of their superordinates and of their peers in the school setting, and this trend probably will become more pronounced in the years ahead. Teachers, for example, are evaluating the work of their principals and of their fellow teachers, and these evaluations are providing the basis for administrative decisions which are of importance to all school personnel.

Theoretical Basis of the Study

Since these evaluations are of such consequence in the school organization, the question of their accuracy is of considerable importance. Scholars, particularly from the discipline of psychology, have hypothesized that certain personal variables of an individual interfere with his ability to perceive and to evaluate another person with absolute accuracy. These hypotheses have been examined in a considerable number of diverse studies carried out in attempts to relate personality traits, as defined in rather broad and unrestrictive terms, of perceivers to inconsistencies or errors in the perception of other people.

Several authors have made rather extensive reviews of these studies in attempts to isolate findings which appear with a degree of consistency. Taft (3) was able to draw a few generalizations based on somewhat nebulous evidence, but his review and other examinations of the research literature indicate a considerable amount of inconsistency in the findings of various studies reported in the journals. This inconsistency was observed by Chance and Meaders who stated that "It . . . seems obvious that the perceiver's personality organization should to some extent affect his perceptions. However, surveys of the research literature . . . and several articles appearing since find only rather fragile support for this 'obvious hypothesis'" (1:200). In another review of the literature in this general area, Zalkind and Costello were able to state only in a spirit of caution that "Research . . . suggests that our perceptions may characteristically be distorted by . . . traits that we possess" (4:226).

Most of the studies reviewed by the aforementioned writers were concerned with accuracy of the perception of personality characteristics of other people, and relatively few dealt with accuracy of the perception of behavior *per se* of other individuals. However, the research reported in the following paragraphs was concerned chiefly with the latter variable because of its consequence to administrative and organizational theory and because of the inconclusive findings of prior studies.

It was assumed that the isolation of relationships between personal variables of individuals and their perception of the behavior of other people would be reason to doubt the ability of a person to evaluate accurately the behavior of another person in the school setting. These hypothesized relationships would indicate that personal variables of the perceiver have a contaminating effect on his perception of the performance of another person in the organization. The purpose of this study was to investigate the existence of certain of these possible relationships.

Instruments Utilized

Three instruments were used during the data gathering phase of the study. The OCDQ was employed to measure perceptions by teachers of the behavior of their principal and the collective behavior of their fellow teachers. This instrument consists of eight subtests defined as Disengagement, Hindrance, Esprit, Intimacy, Aloofness, Production Emphasis, Thrust, and Consideration. As explained by Halpin and Croft in their discussion of the development of the OCDQ, the first four subtests pertain to the teachers' behavior, while the latter four pertain to the principal's behavior (2:29–32).

The MTAI was used to provide a measure of teacher attitude toward children. A low score on this inventory is indicative of a "poor" attitude toward children, while a high score conversely is evidence of a "good" attitude toward the pupils in the classroom. The third instrument employed in the study, the 16 PF (Form C), provided, as its name implies, an evaluation of sixteen different personality variables. Each variable is arranged as a continuum (for example, Factor B has as its poles Dull and Bright), and the score earned by a respondent places him at a spot on the continuum.

Collection of Data

The basic population for the study consisted of the 285 elementary schools located in member districts of the Educational Research and Development Council of the Twin Cities Metropolitan Area, Inc. This number was reduced to 154 elementary schools—forty in Minneapolis, twenty in St. Paul, and ninety-four in the suburban districts—through the establishment of criteria that (a) the principal had to be a full-time supervisor, (b) he had to be employed full time in one particular school, and (c) he had to be serving at least his second consecutive year as principal of this school.

From this population of 154 elementary schools, a stratified random sample of fifty suburban schools and thirty-one Twin Cities schools was drawn without replacement. This procedure resulted in a final sample of eighty-one elementary schools including fifty schools in the suburbs, twenty schools in Minneapolis, and eleven schools in St. Paul.

The instruments were answered by the teachers during a quasi-formal testing session in each of the eighty-one elementary schools included in the sample. Each teacher in a school responded to the OCDQ and, in addition, to either the MTAI or the 16 PF Questionnaire. A method of randomization was used to determine which of the latter two instruments would be answered by each teacher. A few teachers were absent on testing days, one teacher failed to answer the two instruments, and three teachers misunderstood the directions and returned answer sheets which could not be tabulated. However, there was no reason to believe that these few exceptions were systematic and would contaminate the large sample in any way. The final usable sample consisted of 1,691 teachers in eighty-one elementary schools; of these individuals, all answered the OCDQ, 842 completed the MTAI, and 849 responded to the 16 PF Questionnaire.

Examination of Hypotheses

Two major null hypotheses, subdivided into a total of 136 minor null hypotheses, were examined in the investigation of the theoretical concepts under consideration. For the sake of consistency, an alpha of .05 was used in all statistical tests. These hypotheses, the techniques used in their

Table 1.—Correlations Between MTAI Scores and OCDQ Subtest Scores

Teacher Attitude	Dimensions of Climate	Coefficients of Correlation
MTAI	Disengagement	- .1860
MTAI	Hindrance	- .1915
MTAI	Esprit	+ .1802
MTAI	Intimacy	+ .1313
MTAI	Aloofness	- .0799
MTAI	Production Emphasis	- .0592
MTAI	Thrust	+ .1249
MTAI	Consideration	+ .0940

Table 2.—BETA Coefficients Associated with Personality Factors Whose Scores Made a Significant Contribution to the Prediction of OCDQ Subtest Scores

OCDQ Subtests	16 PF Questionnaire Personality Factors A	B	C	E	F	G	H	I	L	M	N	O	Q^1	Q^2	Q^3	Q^4
Disengagement				+.25				-.21	+.28	+.20	+.19					
Hindrance		-.19						-.19		+.16	+.18					
Esprit							+.28	+.19	-.42				-.21			
Intimacy							+.20	+.13								
Aloofness				+.15						+.22						
Production Emphasis																
Thrust							+.24	+.27								
Consideration					+.13		+.15									+.13

examination, and the conclusions concerning their tenability are discussed next.

Null Hypothesis 1. There are no relationships between teacher perception of the dimensions of organizational climate, as measured by the OCDQ, and teacher attitude toward children, as measured by the MTAI.

Eight Pearson Product Moment coefficients of correlation, shown in Table 1, were calculated. Each coefficient revealed the relationship between the MTAI scores and the scores on one dimension of organizational climate.

While all eight r's were quite low, confidence intervals based on a z-transformation technique showed seven of them to be significant at the prescribed level. Of considerable importance was the fact that positive correlations were found between MTAI scores and scores of the four subtests in which high scores are indicative of an open climate, while negative correlations were found between MTAI scores and scores of the four subtests in which high scores are indicative of a closed climate. Since teachers with a "good" attitude toward children tended to perceive all eight dimensions of climate in a manner indicative of an open climate while teachers with a "poor" attitude toward children tended to view all eight dimensions in a manner indicative of a closed climate, the null hypothesis of no relationships was rejected.

Null Hypothesis 2. There are no relationships between teacher perception of the dimensions of organizational climate, as measured by the OCDQ, and personality factors of teachers, as measured by the 16 PF Questionnaire.

Eight multiple regression equations, in which the sixteen personality variable scores served as predictors of the eight OCDQ subtest scores, were calculated. After each of the eight multiple correlation coefficients was determined to be significant at the prescribed level, each of the 128 partial regression coefficients was tested for significance through the employment of an appropriate t-test.

Twenty-two of these beta coefficients, associated with ten of the sixteen personality factors and seven of the eight dimensions of organizational climate, were found to be significant at the prescribed level. An examination of the data, shown in Table 2, revealed that certain personality factors, particularly Factor I (Tough vs. Sensitive), Factor H (Timid vs. Adventurous), and Factor M (Conventional vs. Eccentric), were related to the perception of certain dimensions of organizational climate, especially Disengagement, Hindrance, and Esprit. Of considerable significance is the fact that in no instance did a particular score on a specific personality factor predict one subtest score in a manner indicative of an open climate and a second subtest score in a manner indicative of a closed climate. For example, a high Factor I score was predictive of low Disengagement and Hindrance subtest scores, and it was predictive of high Esprit, Intimacy, and Thrust subtest scores. Because of these various findings, the null hypothesis of no relationships was held to be untenable.

Any interpretation of the findings must be tempered by the fact that (a) many of the minor null hypotheses could not be rejected at the prescribed level, and (b) a number of the relationships found to be significant were quite low and rejection of the null hypotheses resulted from the large sample. However, the data in the two tables reveal that teachers with certain attitudinal and personality traits tended to view one or more of the dimensions of organizational climate in a manner indicative of an open climate or a closed climate. Thus, the research indicates that certain attitudinal and personality variables of teachers are related to their perception of the behavior of their principal and fellow teachers, and these relationships, in turn, indicate error in the process of perception and evaluation.

Implications of the Study

The implications of this study to the administration of schools are of consequence. Before an administrator makes a decision based on an evaluation by certain individuals of the behavior of other persons in the school organization, he should be aware of these basic questions:

> *Have I been informed of the actual behavior of the person or persons under investigation, or have I been informed of behavior as perceived inaccurately by those who reported to me? If the latter is true, to what extent is my information contaminated by errors in perception?*

The research literature reported in various journals and the findings of this study indicate that error probably does occur in the perception of the behavior of other people, but it would be extremely difficult to determine with precision the nature, amount, and direction of the total error. Halpin and Croft, in their description of the development of the OCDQ, recognized this problem when they asked:

> *. . . how much of the description . . . given by a respondent is determined by the actual stimulus-object . . . and how much is determined by elements within the perceiver which operate, as it were, almost independently of the person who presumably is being described? (2:9).*

The existence of perceptual errors and the resultant inaccurate descriptions of the behavior of other people, supported to a degree by the findings of this study, is becoming increasingly more pertinent to the actual administrative process in school organizations. At all levels of education—elementary, junior high, senior high, undergraduate, and graduate—evaluations of the behavior of individuals are occurring through both formal and informal channels, and these evaluations appear to be growing in number, scope, and purpose. As indicated earlier, superordinates long have assumed the function of evaluating their subordinates in the school organization, and now a trend is ever more discernible for those lower in the hierarchy to evaluate those in higher positions of authority.

It is a near certainty that injustices have occurred in the past and could take place in the future because of misinterpretations by school administrators of descriptions of the behavior of people at different hierarchical levels of the organization. These probable injustices will be eliminated or reduced in number only if persons in positions of authority recognize the probability that perceptual errors do occur and that individuals do not perceive with absolute accuracy the behavior of other people in the school organization.

Replications of this study, utilizing different data gathering devices and different samples of school personnel, are a prerequisite to the strengthening of theory in the area of perceptual error. Additional research is needed in order to isolate with a greater degree of certainty the personal variables of respondents which are related to errors in the perception and evaluation of the behavior of other people. Until these relationships are more clearly defined and understood, school administrators should exercise caution when making decisions based wholly or in part on evaluations of the behavior of other people.

REFERENCES

1. Chance, June Elizabeth; Meaders, Wilson, "Needs and Interpersonal Perception," *Journal of Personality,* 28:200–209, June 1960.
2. Halpin, Andrew W.; Croft, Don B., *The Organizational Climate of Schools,* Midwest Administration Center, The University of Chicago, Chicago, 1963, 130 pages.
3. Taft, Ronald, "The Ability To Judge People," *Psychological Bulletin,* 52:1–23, January 1955.
4. Zalkind, Sheldon S.; Costello, Timothy W., "Perception: Some Recent Research and Implications for Administration," *Administrative Science Quarterly,* 7:218–235, September 1962.

I WOULD exclude most white teachers from black ghetto schools because most of them would not be able to pass the . . . ability tests required to teach emotionally and economically deprived black children.—Charles Hamilton, co-author of *Black Power.*

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

The College Environment: Its Meaning to Academically Successful and Unsuccessful Undergraduates[1]

DENNIS EARL NELSON
Prairie View A & M College and American Institutes for Research

ABSTRACT

Forty-two honor roll and forty-two academic probation students responded to eleven concepts relevant to the university environment along nine semantic differential scales representing three major dimensions of meaning (i.e., evaluative, potency, activity). The nature of any differences in ratings by these two groups was sought. A score for each group along each of these dimensions for each concept was derived, resulting in thirty-three t-test comparisons. The evaluative dimension accounted for seven of the fourteen differences found to be significant at $p < .05$ or beyond. The concepts of "Grades," "Myself as a Student," and "Studying" yielded the greatest and most consistent differences. These concepts were rated as less positive, less potent, and less active by unsuccessful students with one exception. "Studying" was viewed as much more active by unsuccessful than by successful students.

"THE SAME word means different things to different people." This statement or one of a similar nature has all too often placed the area of subjective meaning beyond the power of systematic investigation. However, several approaches have been proposed to explore this intriguing area. One such approach has been proposed by Osgood and his associates at the University of Illinois (5). It has become known as the semantic differential (SD).

Numerous studies employing the SD have utilized college students as Ss. For example, Cook (1) attempted to discover semantic predictors of achievement. Husek and Whittrock (2) studied the student ratings of the concept of "teacher" and Stone and Sinnett (8) have sought to learn something about the meaning of the concept "college education" to undergraduates. However, there exists no known research which has explored the meanings of concepts from the college environment among academically successful and unsuccessful college students. The present investigation attempted to determine if differences exist between the semantic differential ratings given such concepts by academically successful students and academically unsuccessful students. It also sought to reveal the nature of any differences uncovered.

The Semantic Differential

Many methods have been utilized to obtain a measurement of meaning. Physiological approaches have sought a correlation of muscular events and thoughts. Salivary reactions and Galvanic Skin Response techniques have been used. Learning methods have included association and scaling devices, among others. Osgood (5) pointed up the weak points of former approaches and proposed a new device, the SD technique, for attacking the problem. Purporting to be a combination of associative and scaling methods, the device grew out of research on synethesia. This is the phenomenon wherein individuals attach sensations in one sense mode to sensations in another sense or mode, for example, the systematic relations of musical notes to colors. Subjects are asked to rate selected words or phrases along 7-point bipolar objective scales. For example, the word "house" might be rated along continua such as "up-down," "black-white," and the like. Intensive cross-cultural and factor analytic work resulted in the establishment of three general factors of meaning which are assessed by the SD technique.

Evaluative is the label given the most dominant factor. Such scales as "good-bad," "beautiful-ugly," and "sweet-sour" demonstrated very high loadings on this factor. Secondly, a potency factor was apparent. This dimension of meaning was represented by such scales as "large-small," "strong-weak," "heavy-light," and others. Scales such as "fast-slow," "active-passive," and "hot-cold" symbolized a third factor which was designated as the activity dimension. These three factors of meaning have held up with amazing consistency during 10 years of research (4).

Hypotheses and Experimental Procedures

1. *The SD ratings of the academically successful students will tend to be higher on those scales representing an evaluative dimension than will the ratings of the academically unsuccessful students.*
2. *There will be no difference in ratings by the two groups on scales representing the potency dimension nor on scales representing the activity dimension.*

Sample. All undergraduates enrolled on a full-time basis at Brigham Young University during the fall semester previous to the execution of this study and appearing on the official university honor roll for that semester formed the academically successful population from which a random sample of one hundred was drawn. Similarly, one hundred students were randomly selected from those officially placed on academic probation by the university during the same semester. A table of random numbers (7) was utilized for both selection procedures.

Materials. A SD booklet containing some brief written directions for the S and a total of ninety-nine items was used. The items were produced by pairing each of eleven concepts with each of nine bipolar adjective scales. Of the nine scales along which each concept was rated, three represented Osgood's evaluative dimension, three represented the activity dimension, and three represented the potency dimension. Each scale was of the 7-position type and the S was instructed to place a "check" mark at the position along each scale that he determined would best represent his present feelings.

Scales Used. The nine scales used were chosen from a study of Osgood's (6) semantic atlas as well as various SD research studies. The evaluative dimension was represented by the scales "good-bad," "beautiful-ugly," and "sweet-sour." "Large-small," "strong-weak," and "heavy-light," were the scales selected to be representative of the potency dimension. Finally, the factor of activity was exemplified by the scales "fast-slow," "active-passive," and "hot-cold." These particular scales appear to possess the highest loadings on the specific dimension which they represent, as well as apparently being the most stable and consistent portrayers of that dimension.

Validity for the instrument thus formulated was established by randomly selecting five Ss from each of the two samples of one hundred and carrying out personal interviews with each of them. The split-halves method, which in this instance meant abstracting and comparing the ratings on the even numbered items with the ratings given on the odd numbered items, was used to determine reliability.

Concepts Rated. Major attempts to measure the environment of higher education were reviewed and several concepts were chosen as a result of this review. These concepts included "college classes," "college regulations," "college teachers," "grades," "roommates," "studying," and "the campus." The concept "L.D.S. Religion" was included due to the fact that Brigham Young University, located in Provo, Utah, is owned and operated by the Church of Jesus Christ of Latter-Day Saints (Mormon). The subculture associated with that religion is a definite part of the institutional environment.

Treatment Procedure. No information was given the Ss either before, during, or after the administration of the device concerning the purpose of the research. All questions concerning S selection were met with the answer that the students had been selected through a random sample. As Ss arrived at the testing center, they were given a booklet and instructed to read the directions carefully. Any questions they had were answered before they began responding to the device. No time limit was set nor were the Ss interrupted or encouraged to change their speed of response. Following two prearranged administration sessions, those Ss in the sample pools who had not yet responded to the device were contacted personally. When efforts to contact Ss were terminated, a total of forty-two honor roll students and forty-two students on academic probation had responded to the booklet. Attrition in Ss, however, was due primarily to two causes. Prior to the study it was found that while eighty-eight of the one hundred honor roll students selected for the study were still in attendance at the university, only fifty-one of the one hundred students on academic probation were still in attendance. Secondly, analysis of Ss contacted after the completion of the investigation revealed that those students with whom it was impossible to find a mutually agreeable time for administration of the SD were overwhelmingly those from the honor roll group.

Scoring. For scoring purposes, each position along the 7-position scale was given a number, commencing with one, designating the position at the far left, and ending with seven, designating the position on the far right of the scale. If the S's check was placed in the one or two position at the left or the six or seven position at the right, the meaning of that concept for the S would be quite closely related to the adjective placed on that side of the scale in question; if placed in the number three or five positions, slightly related; and if in position number four, either equally related or entirely irrelevant in meaning as far as that particular individual was concerned.

To determine a particular individual's score on a given concept, an average was obtained by summing the position numbers marked along each of the three scales representing a particular factor of meaning and subsequently dividing by the number of scales along which the concept had been rated (e.g., if an individual marked position four along the "good-bad" scale, position two along the "beautiful-ugly" scale, and position three along the "sweet-sour" scale, the average score for this S on the evaluative scales for this concept would be three). This would indicate that for this S, the concept carried a slightly positive meaning, evaluatively speaking. To obtain scores for the two groups on the three factors of meaning for each concept, these raw position scores of individuals on the relevant scales were summed and averaged. These procedures yielded thirty-three scores for each group (i.e., a group score on each of three factors of meaning for each of eleven concepts).

Data Analysis. Specific support for the acceptance or rejection of hypotheses was obtained through t-test analysis procedures (3). A total of thirty-three t-tests was calculated to determine if significant differences existed on any of the pairs of group scores obtained. The results of these analyses are reported in Table 1.

Table 1.—Results of t-Test Procedure

Variable	T-Ratio	Sig
L.D.S. Religion—E	-2.885	.01
L.D.S. Religion—P	-0.550	NS
L.D.S. Religion—A	-0.668	NS
College Teachers—E	-1.786	NS
College Teachers—P	-0.323	NS
College Teachers—A	-1.371	NS
Grades—E	-5.138	.001
Grades—P	-4.239	.001
Grades—A	-2.877	.01
Roommates—E	-1.004	NS
Roommates—P	1.307	NS
Roommates—A	-0.883	NS
Myself as a Student—E	-3.137	.01
Myself as a Student—P	-2.961	.01
Myself as a Student—A	-4.699	.001
The Ideal Student—E	-1.420	NS
The Ideal Student—P	0.266	NS
The Ideal Student—A	-2.351	.05
Studying—E	-3.093	.01
Studying—P	-2.910	.01
Studying—A	2.220	.05
The Campus—E	-2.753	.01
The Campus—P	-1.617	NS
The Campus—A	-0.734	NS
College Regulations—E	-0.104	NS
College Regulations—P	-0.523	NS
College Regulations—A	-0.996	NS
College Classes—E	-2.654	.01
College Classes—P	-0.349	NS
College Classes—A	-1.737	NS
A College Education—E	-2.191	.05
A College Education—P	-1.016	NS
A College Education—A	-0.996	NS

T limits for significance at 80 degrees of freedom: .05 level = 1.990; .01 level = 2.638.

T limits for significance at 80 degrees of freedom: .001 level = 3.431.

This t limit is an interpolation resulting from the fact that only columns for 60 and 120 degrees of freedom were available in the statistical text used.

Results and Discussion

The major concern of the study as reflected in the hypotheses stated earlier was to determine whether or not ratings of selected concepts by each of two student groups (i.e., academically successful and unsuccessful) along three dimensions of meaning for each of eleven college environment related concepts would differ. Hypothesis 1 received considerable support from the evidence presented in Table 1. Through comparison of the mean ratings of the two groups given the concepts along the evaluative dimension, all eleven were viewed more favorably by the academically successful students and in the case of seven of the eleven concepts rated, the difference was statistically significant. Some evidence was also gleaned which would call into question the second hypothesis, stated as it is in the null form. An examination of group means along the potency dimension of meaning revealed that there were indeed some differences between the two groups. The mean ratings given by the honor roll students on this dimension for every concept appeared to be somewhat higher than the corresponding ratings of the academically unsuccessful students, and in three cases the differences proved to be significant. One finds rather consistent differences of a like nature in mean ratings given in the activity dimension and with regard to four concepts the differences reached statistical significance.

It was striking to note that seven of the fourteen instances in which the ratings were found to differ significantly occurred in relation to the evaluative dimension. Thus, the connotative meaning of the college environment seemed to differ most in this positive versus negative or good versus bad realm of meaning. The concept of "grades" produced the greatest difference in the ratings provided by the two groups. The academic probation students consistently saw grades as a less positive entity than those on the honor roll. Similarly, they viewed the concept as less potent or powerful and less active or dynamic. It may be postulated that these meanings are a result of not being successful in attaining desirable grades. Such individuals may in general see grades as therefore having less value, less influence or importance, and as a more passive element in college life. The differences on this concept were so marked that they should be investigated further. Just as "grades" was the characteristic used to formulate the two groups of Ss, so "grades" highlighted the differences in the meaning the college environment held for these students.

A second concept yielding significant differences on all three dimensions of meaning was "myself as a student." The magnitude and consistency of the differences revealed here may indicate the relevance of further investigation aimed at just how self-concept affects an individual's behavior. Those on academic probation saw themselves in a less positive light and as being less dynamic and active than did honor roll students.

"Studying" was the third concept yielding substantial differences. While those students on academic probation viewed studying less favorably and as being less powerful or potent, they saw it as being significantly more active a phenomenon than did the honor roll students. Thus, the concept of "studying" was associated by the academically unsuccessful student more closely with adjectives such as "fast," "active," and "hot," whereas the associations reflected in ratings by the honor roll students tended more toward adjectives such as "cold," "passive," and "slow." On the surface, this result seemed rather opposite from what one might logically expect. One plausible explanation for this data suggests itself if one remembers that, as a group, the academic probation students saw themselves as less active and dynamic than they viewed "the ideal student" as being. Similarly, their ratings of themselves and grades on the same dimension were quite divergent. Keeping in mind these findings, the association of "studying" with adjectives such as "fast," "active," and "hot" by academic probation students seems less unique. Indeed, it seems reasonable that such students would view "studying" as being associated with adjectives similar to those associated with "grades" and "the ideal student" and rather different from those adjectives associated with themselves.

In addition to the consistent differences on these three concepts, there were isolated instances of differences on three other concepts. On one dimension of meaning, these groups differed in their view of "the ideal student." On the activity dimension, honor roll students tended to rate "the ideal student" as being more closely associated with adjectives such as "fast," "active," and "hot" than did the students on academic probation. Additionally, the academic probation Ss possessed less favorable meanings of the concepts of "the campus" and "L.D.S. Religion." These differences might suggest that their unsuccessful experiences academically were generalizing to some degree and affecting how they felt toward the campus in general and toward the religious atmosphere at the school. Other feasible interpretations of the data, however, could be generated. Studies on the religious activity of students on academic probation and of the degree to which aggressive acts against campus facilities are associated with students who are on academic probation provide two examples of the kind of data that would be useful in understanding the possible ramifications of these differences.

The differences in how the two groups rated themselves as students and "the ideal student" should be noted. Both groups found more than a modicum of agreement on the meaning of "the ideal student" but differed widely on their ratings of themselves as students. While it is not possible to determine from the present study whether these perceptions were results of the college experience or whether such meanings were held by the Ss prior to entering college, research aimed at revealing to what degree unfavorable or favorable experiences in the college setting contribute to the connotative meanings manifest in the ratings of these groups of students is to be encouraged. Studies offering a far greater number of concepts from the college environment should be undertaken to uncover other differences between the two groups.

The research, herein described, admittedly is exploratory in nature and raises several additional questions for future investigation. In addition to those avenues of research already suggested, studies employing students who are members of particular subcultures within our society (e.g., Mexican–American, members of a distinct religious denomination) should be undertaken to discover if there is indeed a psychological "frame of reference" which affects the connotative meanings such individuals give to the college environment.

FOOTNOTE

1. The assistance of the following is acknowledged in the execution of this research: Research Assistants Janet Forbes, Alyce Ford, Harmon Jensen, and Boyd Nelson, and Brigham Young University, Office of Academic Standards.

REFERENCES

1. Cook, D. R., "The Relationship of Concept Meaning to Achievement," *Psychological Researches*, 6:118–123, 1962.
2. Husek, T. R.; Whittrock, M. C., "The Dimensions of Attitudes Toward Teachers as Measured by the Semantic Differential," *Journal of Educational Psychology*, 53: (no. 5) 209–213, 1962.
3. McNemar, Q., *Psychological Statistics*, John Wiley and Sons, Inc., New York, 1962.
4. Miron, M. S.; Osgood, C. E., "Language Behavior: The Multivariate Structure of Qualification," in Cattell, R. B. (ed.), *Handbook of Multivariate Experimental Psychology*, Rand McNally, Chicago, 1966.
5. Osgood, C. E., "The Nature and Measurement of Meaning," *Psychological Bulletin*, 49: (no. 3) 197–238, May 1952.
6. Osgood, C. E.; Suci, G. J.; Tannenbaum, P. H., *The Measurement of Meaning*, University of Illinois Press, Urbana, 1957.
7. Peatman, J. G., *Descriptive—Sampling Statistics*, Harper and Brothers, New York, 1947.
8. Stone, L. A.; Sinnett, E. R., "The Meaning of a College Education as Revealed by the Semantic Differential," *Journal of Counseling Psychology*, 11: (no. 2) 168–172, 1964.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Assessment of Efforts to Aid Socially Isolated Elementary School Pupils[1]

MERL E. BONNEY, North Texas State University

ABSTRACT

This study was focused on an assessment of special efforts to aid pupils in the lowest fourth of social status with their classmates. During one semester in four classes the effects of seventeen different kinds of socializing experiences were assessed as compared with social status scores the preceding term and as compared with similar data collected in two control schools. All group differences were statistically insignificant, but a few individuals made marked gains. These largely negative findings are explained on the bases of concepts from sociometrics, social psychology, Gestalt principles, and the psychology of perception. School personnel are warned against assuming that most socially low pupils can be aided by socializing experiences.

PROBABLY ALL students in college classes bearing on child development and classroom management have been given suggestions by their teachers and textbooks on what to do to help socially isolated pupils. Most of these suggestions focus on ways whereby such a pupil can be brought into more interpersonal relationships with his classmates, or how he can make some kind of contribution to his group. The assumptions involved in these suggestions are that if an isolated child were better known he would be better accepted, and if he could make a contribution to his group he would be better appreciated. Both of these assumptions have certain face validity but many teachers who have acted upon the suggestions mentioned will testify that for reasons not very clear to them their efforts have frequently been ineffective. Yet it seems certain that this frequent ineffectiveness is not sufficiently recognized by most college teachers and textbook writers who deal with the topic of socially isolated pupils.

During the academic year of 1967–68, the writer conducted an investigation of the extent to which socially weak elementary school children could be helped by calculated efforts to do so, with the expectation that such efforts would prove to be quite effective with most of the Ss.

Subjects and Measurements

Subjects for this research were pupils in grades 3 through 6 in three schools. One of these schools was the North Texas State University Laboratory School, another was a public school in Denton, Texas, and the third was a private school in Dallas, Texas. All pupils attending these schools responded to two measurements in September, January, and May. The first measurement was a sociometric test which asked each child to list the names of other pupils in his class whom he considered to be (a) the best leaders—ones who could get things done and influence others, and (b) the ones he would prefer as associates on a classroom work assignment, and (c) the ones he would prefer in a play activity. A child's score on each of these criterion was the total choices received, with no weighting utilized.

The second measurement was a social roles instrument consisting of eighteen items bearing on different kinds of social performances which children on these grade levels engage in, such as making an oral report, helping settle a group conflict, taking part in a dramatization, and offering good suggestions in class discussions. The pupils wrote under each item the names of others in their class whom they considered to be "one of the best" in this particular kind of group functioning. This instrument was devised to get a frequency count on the extent to which each child was perceived by his classmates as being "good at" the eighteen social roles. No negative or unfavorable roles were listed; consequently, the higher a child's score, the more he was perceived as performing well in a wide variety of socially approved classroom social roles. From the three sociometric criteria and the social roles measurement, certain pupils in each class in all three schools were placed in a *low* or socially weak category. In most instances these pupils were in the lowest fourth of their respective classes on three of the four assessments listed. A few were low on all four, and all of them were in the lowest fourth on at least two of the four distributions.

The number of children in the twelve classes utilized in this study varied from twenty to thirty, with the most usual number being twenty-five.

The only *low* pupils who were the objects of special efforts to help them improve their social status with their classmates were those in the four grades in the Laboratory School, and these efforts were confined to the second semester. The classes in the Denton and Dallas schools were utilized as comparable groups in which no special efforts were instigated.

The question may be raised as to whether the pupils in the *low* categories in the Laboratory School were inferior to the rest of their respective classes in assessments other than the sociometric and social roles. In regard to IQ's obtained from the California Test of Mental Maturity, there were no reliable differences, but in reference to Stanford Achievement Test scores the means for the "rest of the class" were in all four classes higher than for the lows, and in two grades these differences were statistically reliable. No doubt this generally poor record of academic performance on the part of the lows should be viewed both as a cause and a consequence of their inferior social functioning.

The lows in the Laboratory School consisted of approximately 60 percent boys and 40 percent girls. This was true at all three testing times.

The data on IQ's, academic achievement, and sex distribution for the lows in the Denton public school and the Dallas school were quite similar to those obtained in the Laboratory School.

Procedures and Socialization Efforts

The special efforts to help the lows in the Laboratory School took place between the second week in February and the second week in May. These efforts were conducted by the writer and his research assistant. They worked in close collaboration with the four classroom teachers, the special teachers of art, music, and physical education, and the school librarian. Each teacher had a list of the seven pupils in each class who were to be helped.

Records on what was done to aid the low children were obtained (1) from direct observations in the classrooms by the writer and his assistant —to the extent of approximately 6 hours a week combined time and, (2) from conferences two or three times a week with the teachers on what they had done or had observed.

In addition to numerous suggestions talked over with the teachers during the conferences, four "newsletters" were distributed to all the participating Laboratory School teachers during the spring semester. These consisted of two or three mimeographed pages of brief descriptions of things done in the Laboratory School which were considered to be good examples of the kind of teacher behavior, or group-initiated projects, of value in promoting socialization objectives. Some of these were endeavors in which the lows participated along with the rest of the class, but most of them were descriptions of things done to aid low pupils. Examples of the latter are:

> *1. A girl who is capable in a particular physical skill was asked to demonstrate how to do this skill to the rest of the class.*
>
> *2. One teacher holds a weekly conference with the mother of a child to help him in a program of developing more responsibility at home.*
>
> *3. A child was asked to take home a story* (*that the class had not heard)* and to prepare it well *for reading or telling to the class.*
>
> *4. One teacher makes sure that a girl who is over-identified with boys is always placed in a group composed of both boys and girls.*

All of the items in the newsletters were of a positive nature—things considered desirable and therefore worthy of emulations and extensions. No teacher's name was mentioned. The purpose of these newsletters was to encourage the teachers to initiate in their respective classes things similar to those being done by other teachers in their building. That this purpose was at least partially achieved was indicated by remarks from the teachers during the term, and also by their unsigned evaluative comments made at the end of the semester. The idea of behavior change involved here is closely akin to the equilibrium theory tested in a teacher population by Gage, Runkel, and Chatterjee (10).

More specifically, what did the records from the observations and teacher reports show in regard to the kinds of individual behaviors or participation experiences which would presumably aid the social status of the low pupils? These data are given in Table 1.

Results

Sociometric Choices. In assessing changes in sociometric choices, the first step was to convert all raw scores to percentages since the number of choices any child could receive was certain to be influenced to some extent by the total volume of choices expended by his group.

The following points were the focus of attention: (1) the percentage gains or losses of each low child in the four Laboratory School classes between September and January, on the three sociometric questions combined, as compared with the gains and losses of each low child between January and May, and (2) the comparisons of these gains and losses with comparable data obtained on the low pupils in each of the four grades of the other two schools.

Table 1. Frequencies of Various Procedures and Experiences Assumed to be of Value in Aiding Sociometrically Low Pupils to Attain Higher Choice Status

Procedures and Experiences	Third* Grade	Fourth Grade	Fifth Grade	Sixth Grade
1. Collaborative efforts with two or more other pupils in free activity period—solving a puzzle, making a bulletin, playing a game, or in a work group project	13	14	16	11
2. Interview with a parent about child's adjustment in school and how the parent could help	14	20	14	10
3. Participated well, or fairly well, in class dramatization or role playing in classroom or auditorium	12	12	11	7
4. Made a creditable showing in some performance before the class, as in art, music, physical education, or in oral report	19	20	9	10
5. Brought some kind of object, pet, or book to class and shared it with other pupils	8	6	7	6
6. Teacher exerted influence with child or parent to get child in a community group, or to invite other children into the home	7	5	2	0
7. Special class assignment by teacher—passing out materials, helping another child, doing a room chore, being chairman of a group	10	12	9	9
8. Successful part in talent show or hobby show	2	5	2	10
9. Commended by teacher before class for exceptional test score, or other form of academic superiority, or name put on bulletin board	9	7	3	12
10. Chosen by pupils for a minor class office—row captain, secretary, chairman of committee, or one to give a report	9	6	5	8
11. Seated near or placed at same table with others he chose—sometimes with others of high choice-status	8	7	9	5
12. Gave an acceptable suggestion in a class discussion	17	18	12	16
13. Put on a timed schedule to help finish work on time—keeping chart, records, holding to a deadline	0	4	1	3
14. Teacher conference with child about his behavior or work	8	10	9	7
15. Involved in a cooperative, friendly, activity with one other child—not assigned work	11	10	7	16
16. Class discussion of particular kinds of personal-social problems (initiated by teacher)	3	2	4	4
17. Went to individual play therapy or counseling session	30	22	6	0
		(3 pupils)	(2 boys)	(1 boy)

* N=7 for each grade.

When the data for the Laboratory School were assembled, it was immediately evident that practically all the percentage differences for individual children between September and January and between January and May were of a zero order. Between September and January data on percentage changes were complete on twenty-six low pupils in all four grades. Of these pupils, sixteen showed changes in which the first two figures were 00. The others varied from 1 to 3 percent and with three of the scores being in a negative direction.

The results in the two other schools were very similar. In the Denton school data were complete on twenty-seven low pupils in all four grades for the September to January measurements and likewise for the January to May assessments. More than half of the percent scores for changes for both time intervals were of a zero order. All the other changes were 1 or 2 percent, with three of these being minus values.

In the Dallas school, the test-retest data were available on twenty-one low pupils for September to January and likewise from January to May. Over a third of the change scores were zero; the others varied from 1 to 3 percent, with eight of them being in a negative direction.

As would be expected from the foregoing report of findings for individuals, when these data were worked up on a group basis for each class in each of the three schools, practically all of the group percent change scores for both time intervals were of a zero order. Consequently, no statistical analysis on reliability of differences between experimental and control groups was called for.

Mutual Choices. What was found in reference to changes in number of mutually reciprocated choices? These choices were included in this analysis because of the assumption that an individual enjoys greater social satisfactions in a group if the particular persons he prefers to associate with also desire to associate with him.

In calculating mutuals, only the second and third sociometric questions bearing on play and work choices were used, since the first question bearing on leadership did not involve any interpersonal association.

In the Laboratory School, between September and January, the lows as a group gained four mutuals, but between January and May, in two of the classes the change-scores were 0, and in the other two classes these scores were —5 and —2. Obviously the lows lost ground the second semester in establishing reciprocated mutual relationships.

In the Denton public school the lows had fewer mutualities in January than in September. In May they registered no further losses but neither did they improve their relative positions with the rest of their classes on mutualities.

In the Dallas school between September and January the lows gained a few mutuals, but between January and May in three of the classes the lows lost from two to six mutualities, while in the other one there was no change.

Interpersonal Responsiveness of the Lows. Since, as implied before, the number of mutuals any child could establish would obviously depend to some extent upon how many choices he gave, it is pertinent to ask if the lows gave choices in proportion to their representation in their respective classes. In reference to the Laboratory School, the answer is definitely in the affirmative. In only one testing out of the twelve (for the school year), involved in the four classes, did the choosing responsiveness of the lows fall slightly below their proportional representation in their respective classes. In all four grades, the proportions of the total class choices made by the lows increased in May over January. Consequently, it may be concluded that at no time during the school year could the poor social standing of the lows in the Laboratory School be even partly explained on the basis of their lack of responsiveness to others. The old statement "If you would have friends, be one," has never been more than a half-truth.

In reference to the proportional choosing of the lows in the other two schools, essentially the same picture was presented as found in the Laboratory School. Out of the twenty-four figures available on this point for the various testings, only three showed slightly lower proportions than the representations of the lows in their respective populations.

Nominations of Lows for Social Roles. How did the lows in the Laboratory School fare in being nominated "as one of the best" in one or more of the eighteen approved social roles listed on the form presented each class? The most pertinent question for this report is: Did the lows gain in percentage of the total class nominations going to them in the second semester? The answer is that two classes showed small gains of less than 7 percent while the other two showed small losses of approximately 3 percent. The average percentage of total class scores received by the lows in the four classes in September was 10; in January it was 10; and in May it was 11. It will be recalled that the lows constituted approximately one-fourth of their respective classes.

Data on social roles nominations from the Denton school and the Dallas school were very similar to those for the Laboratory School, since all gains for the lows were 6 percent or less, and about half of the scores were either zero or negative.

Exceptional Cases. In spite of the very generally small or negative results from this study there were two or three pupils in each of the three schools under investigation who made exceptional gains on either the sociometric instruments or the social roles nominations, especially between the January and May testings. These exceptional cases underscore a point of general significance from the research, namely, that in future studies of this nature much more focus should be placed on individuals as opposed to concentrating exclusively on group data.

Attitude Toward School Scale. As one aspect of this research, an Attitude Toward School Scale was developed. This scale, consisting of forty-seven items, was composed partly of items taken from other similar scales, partly from classroom observations, and partly from an analysis of the questionnaire responses obtained from 120 fifth and sixth grade pupils who stated how they felt about various aspects of their school situation. When this scale was scored for different degrees of positive and negative attitudes toward the school, the differences between the scores for the lows in three of the four Laboratory Schoo classes and the remainder of their respective classes was 3 points, but this was not statistically significant.

In none of the eight classes in the Denton school and the Dallas school was the difference between the scores for the lows and the remainder of their respective classes statistically significant. These findings with the Attitude Toward School Scale warrant the conclusion that the persistently poor social status of the sociometrically low pupils in all three schools could not be said to be due to their holding indifferent or antagonistic attitudes toward their school situations.

Review of the Research

A review of other studies in the area of this research reveals findings very much in line with those herein reported. Although some investigators such as Amidon (1), Amidon and Hoffman (2), Chennault (4), Cox (5), Dineen and Garr (6), Early (7), Flanders and Havumaki (8) Kranzler, Mayer, Dyer, and Munger (12) have reported some social status gains for children who have been given various kinds of persona social treatments, these results can be criticiz on one or more counts: (1) the gains were qu small, (2) the gains may easily have been of short duration, (3) the number of Ss was very small, (4) the gains were obtained under especially arranged conditions as contrasted with classroom settings, (5) the results were confounded by conditions other than the one under investigation, and (6) the gains in a particular study were inconsistent in reference to other aspects of the collected data.

In their report of a study on the effects of individual counseling on pupils' sociometric choice status in classroom groups, Mayer, Kranzler, and Matthes (13) state that research in this area has generally found no significant improvements from such efforts. In their study, which utilized both individual and group counseling, no significant improvements were shown.

Evidence supporting these generally negative results is found in a report by Bonney and Nicholson (3) on the effectiveness of preschool experiences on subsequent social adjustments in the elementary school grades. Although there were a few exceptions to the general trend, the overall picture (with over 200 Ss) was one of no significant differences in sociometric choices received, or in teacher ratings, between those pupils who had, and those who had not, attended nursery school or kindergarten. Although no special kinds of treatment were tested in this study, the findings do bear on the ineffectiveness of socializing experiences in helping children overcome personality handicaps.

Discussion

The findings of this research, and the results of other studies reviewed above, strongly suggest that much of the advice given teachers and prospective teachers on how to help socially maladjusted pupils is naive and based on some fallacious assumptions. The chief fallacy involved is the assumption that a socially low child's contribution will be objectively evaluated on its merits. This is a fallacy for the following reasons:

> *(1) The responses of the members of a group toward any one individual in it are heavily determined by the expressed attitudes of high prestige members (or the significant others). This is a well established principle in sociometrics and social psychology.*
>
> *(2) The evaluation of anyone's ability performance in a group is very much affected by how this individual is regarded as a person. Interpersonal assessments color all other judgments.*
>
> *(3) Socialization is a two-edged sword—it cuts both positively and negatively. The more all the members of a group respond to each other, the greater are the probabilities of stimulating both positive* and negative *attitudes toward particular individuals. John French (9) demonstrated as early as 1941 that cohesive groups, as contrasted with loosely organized ones, are characterized not only by more we-feeling and interdependence but also by more frustrations and interpersonal aggressions. In other words, the more some low pupils interacted with the rest of their classmates the more they aroused toward themselves both positive* and negative *evaluations. To know some people better is to dislike them more.*
>
> *(4) The worth of any contribution to a group is never determined on a specific basis but always in reference to the contributions of others. In Gestalt terms, the quality of a figure is strongly affected by the ground on which it occurs.*
>
> *(5) How the performance of any individual is perceived is heavily predetermined by a generalized perception of how this person is expected to perform. Much data from the psychology of perception supports this statement.*

When a child of high social status in one of the Laboratory School classes performed, or made any other kind of contribution, he had everything going for him. Other pupils of high prestige responded to him favorably, his personal acceptability accentuated favorable responses throughout the class, usually he behaved so that additional social exposures enhanced his personality image, generally what he had to offer was of high quality within his group, and he performed in a social atmosphere of positive expectations.

By contrast, when a child of low social status performed or made other contributions to his group he frequently had everything going against him. Pupils of high prestige would sometimes ignore or obviously reject his contribution, his poor personal acceptability throughout his group cast a negative halo over his efforts, sometimes he behaved in such ways that his greater visibility worked against him, quite often what he had to offer was below the quality-standard of his class, and he performed against a social atmosphere of derogatory expectations. This latter point was especially strong in the Laboratory School since most of these children had been together for several years, or longer, and consequently had formed very stable social structures and very persistent expectations of one another. Consequently the contributions of pupils of high prestige and of high-choice status were sometimes overvalued whereas the contributions of pupils of low prestige and low-choice status were frequently devalued.

In addition to these conceptual explanations from sociometrics, social psychology, Gestalt principles, and the psychology of perception, it seems very likely that self-concept principles were also operating to help account for the largely negative findings of this research, since it is highly probable that many of the low children developed negative self-images and that these images persistently undermined their capacities to present themselves favorably. It has been well established that many people introject predominant social definitions held in reference to them even when these definitions are derogatory to them.

Finally, some of the lows had personality problems which are very difficult to alter, such as: effeminacy in a boy, overly masculine identifications in a girl, persistent evasion of work responsibilities, marked inability to succeed on the playground, frequent display of bizzare and disturbing mannerisms, and too rigid adherence to adult approved social and moral definitions. It is well known that significant modification of such problems requires intensive efforts over long periods of time with individuals, teachers, and parents.

Space in this article does not permit a detailed consideration of what schools should do to promote the personal-social development of children, but it seems likely that the kind of models which will prove most effective are those provided by Slavson's activity group therapy (15,16), Hobbs' (11) laboratory type schools, and other similar, intensive, holistic programs.

FOOTNOTES

1. This study was made possible through the generous support of funds from Faculty Research of North Texas State University.
2. Details on the test–retest reliability, internal consistency, and scoring procedure for this scale may be obtained from the author at North Texas State University.

REFERENCES

1. Amidon, E., "The Isolate in Children's Groups," *Journal of Teacher Education,* 12:412–416, 1961.
2. Amidon, E.; Hoffman, C. B., "Can Teachers Help the Socially Rejected?" *The Elementary School Journal,* 66:(no. 3)149–154, 1965.
3. Bonney, M. E.; Nicholson, Ertie L., "Comparative Social Adjustments of Elementary School Pupils with and without Preschool Training," *Child Development,* 29:125–133, 1958.
4. Chennault, Madelyn, "Improving the Social Acceptance of Unpopular, Educable, Mentally Retarded Pupils in Special Classes," *American Journal of Mental Deficiency,* 72:(no. 3)455–458, 1967.
5. Cox, F. N., "Sociometric Status and Individual Adjustment Before and After Play Therapy," *Journal of Abnormal and Social Psychology,* 48:354–356, 1953.
6. Dineen, Mary A.; Garry, R., "Effect of Sociometric Seating on a Classroom Cleavage," *Elementary School Journal,* 56:358–362, 1956.
7. Early, Joan C., "Attitude Learning in Children," *Journal of Educational Psychology,* 59:(no. 3)176–180, 1968.
8. Flanders, N. A.; Havumaki, S., "The Effect of Teacher–Pupil Contacts Involving Praise on the Sociometric Choices of Students," *Journal of Educational Psychology,* 51:65–68, 1950.
9. French, J. R. P., "The Disruption and Cohesion of Groups," *Journal of Abnormal and Social Psychology,* 36:361–377, 1941.
10. Gage, N. L.; Runkel, P. J.; Chatterjee, B. B., "Changing Teacher Behavior Through Feedback from Pupils: An Application of Equilibrium Theory," in Charters, W. W., Jr. (ed.), *Readings in the Social Psychology of Education,* Allyn and Bacon, Boston, 1963.
11. Hobbs, N., "Helping Disturbed Children: Psychological and Ecological Strategies," *American Psychologist,* 21:1105–1115, 1966.
12. Kranzler, G. D.; Mayer, G. R.; Dyer, C. O.; Munger, P. F., "Counseling with Elementary School Children: An Experimental Study," *The Personnel and Guidance Journal,* 44:944–949, 1966.
13. Mayer, G. R.; Kranzler, G. D.; Matthes, W. A., "Elementary School Counseling and Peer Relations," *The Personnel and Guidance Journal,* 46:(no. 4)360–365, 1967.
14. Shaw, M. C.; Wursten, Rosemary, "Research on Group Procedures in Schools: A Review of the Literature," *Personnel and Guidance Journal,* 44:27–33, 1965.
15. Slavson, S. R., *The Practice of Group Therapy,* International University Press, New York, 1947.
16. Slavson, S. R., *An Introduction to Group Therapy,* International University Press, New York, 1954.

RESEARCH BRIEFS—Continued from page 344

the total spending. Interestingly, one of the most frequently cited candy rewards, the small chocolate covered by a sugar shell, accounted for only 2 percent of total token spending.

The high density spending results certainly raise a question of the practicality of incorporating a wide variety of rewards in token economies as suggested by Ayllon and Azrin, "*in spite of any inconvenience or administrative difficulty that might result from doing so.*" (1:81)

With regard to spending tokens for high value rewards, two interesting findings were uncovered: girls were more apt than boys to be token savers and high value spenders; white students were more apt to be savers and high value spenders than were Negro students. Future research will further probe the dimensions of student reward choice.

Harold R. Strang, University of Virginia.

FOOTNOTE

1. This research was supported by the Department of Education, University of Virginia, and by Grant OEG 3–70–0007 (010) from the U. S. Office of Education.

REFERENCES

1. Ayllon, T.; Azrin, N., *The Token Economy: A Motivational System for Therapy and Rehabilitation,* Appleton, Century, Crofts, New York, 1968.
2. Strang, H. R., *An Automated Approach to Remedial Reading,* paper presented at American Educational Research Association, New York, February 1971.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

Effect of the System Approach on Changing Preservice Teacher Attitudes Toward Selected Instructional Design Factors

THOMAS G. KENNEDY and ROBERT A. HUMPHREY

ABSTRACT

The purpose of the study was to determine the effect of system approach training upon changing preservice teachers' attitudes toward selected design factors. An experimental and control group posttest design was utilized. The median test was used to determine if significant differences existed between the two groups. Differences were found between the samples in six of the eleven factors studied. A significant difference (.001) was also found between the groups in their overall evaluation of the program.

THE NEED for formulating and implementing a new paradigm for the education of preservice teachers is one of the most pressing problems in education today. Preservice programs must prepare teachers to meet the present as well as the future needs of education in order to be successful. Carl J. Lange, in the report for the governor's conference on education for the future, explains this need in the following statement:

> *Prevalent themes in current discussion of education inevitably include consideration of the unrest and ferment relating to teacher education. There are facets to the general issues of educating teachers. Some of the problems most frequently discussed are: (1) the need for new roles for teachers in order to make effective use of advances in educational technology and new insights into learning; (2) inadequacy of present-day teachers for work with disadvantaged students; (3) improper use of newly developed curricula by teachers; (4) the need for effective utilization of behavioral science by teachers in the instructional process; and (5) a tendency for some professional teacher education courses to be trivial, superficial, or irrelevant* (3).

The implication of the lack of relevancy between existing curricula practices and real-world objectives demands the total review of established educational programs. Merrill, Baird, and Bauer state that "teachers of the future must be prepared to cope creatively with new educational problems and base their professional decisions on theory rather than the folklore of conventional practice" (4). One method that might overcome some of the weaknesses of present educational programs is the systems approach applied to teacher training.

Learning, for this paper, will be defined as a change of behavior resulting from an instructional process or program. The purpose, then, for an instructional design is to develop an instructional *system* for predictable learning.

A system has been defined as "the sum total of separate parts working independently and in interaction to achieve previously specified objectives" (2). Incorporating this definition of a system, the present writers have taken the position that it is necessary to measure and determine the relevancy between all facets of the instructional system to insure a predictable outcome or a measurable product at the end of the instruction period. Therefore, to determine relevancy, the instructional objectives and methods should be based on identified needs. Kaufman states that

> *a need may simply be defined as the difference between what is and what is required . . . a definition that indicates that a need is a measurable difference or distance between a present state or condition and what is required to be accomplished*" (1).

Following the Kaufman definition for identifying needs, the identification of need at the colleg preservice teacher level was determined to be a follows.

What is: Existing teacher education program usually (1) do not instruct preservice teachers i the strategy of educational problem solving (2) do not instruct preservice teachers in desig ing measurable learning outcomes; (3) do n allow preservice teachers to progress at their ow pace; (4) do not instruct preservice teachers design individual instructional material; (5) do not allow validation of instructional design at the student teaching level; and (6) do not instruct preservice teachers in an individualized instruction situation—the educational design of the future.

What is required: Teacher education programs need (1) an analysis of societal, professional, and individual performance specifications necessary for the professional teacher to meet present and projected requirements; (2) a continuous progress curriculum based upon relevant predetermined performance objectives integrated with the student teaching experience; (3) a program that would allow preservice teachers to progress at their own rate; (4) give the student experience as a participant in an individualized program at the preservice teaching level; and (5) instruct the student in designing and implementing instructional materials for individuals based upon predictable and measurable learner terminal behavior.

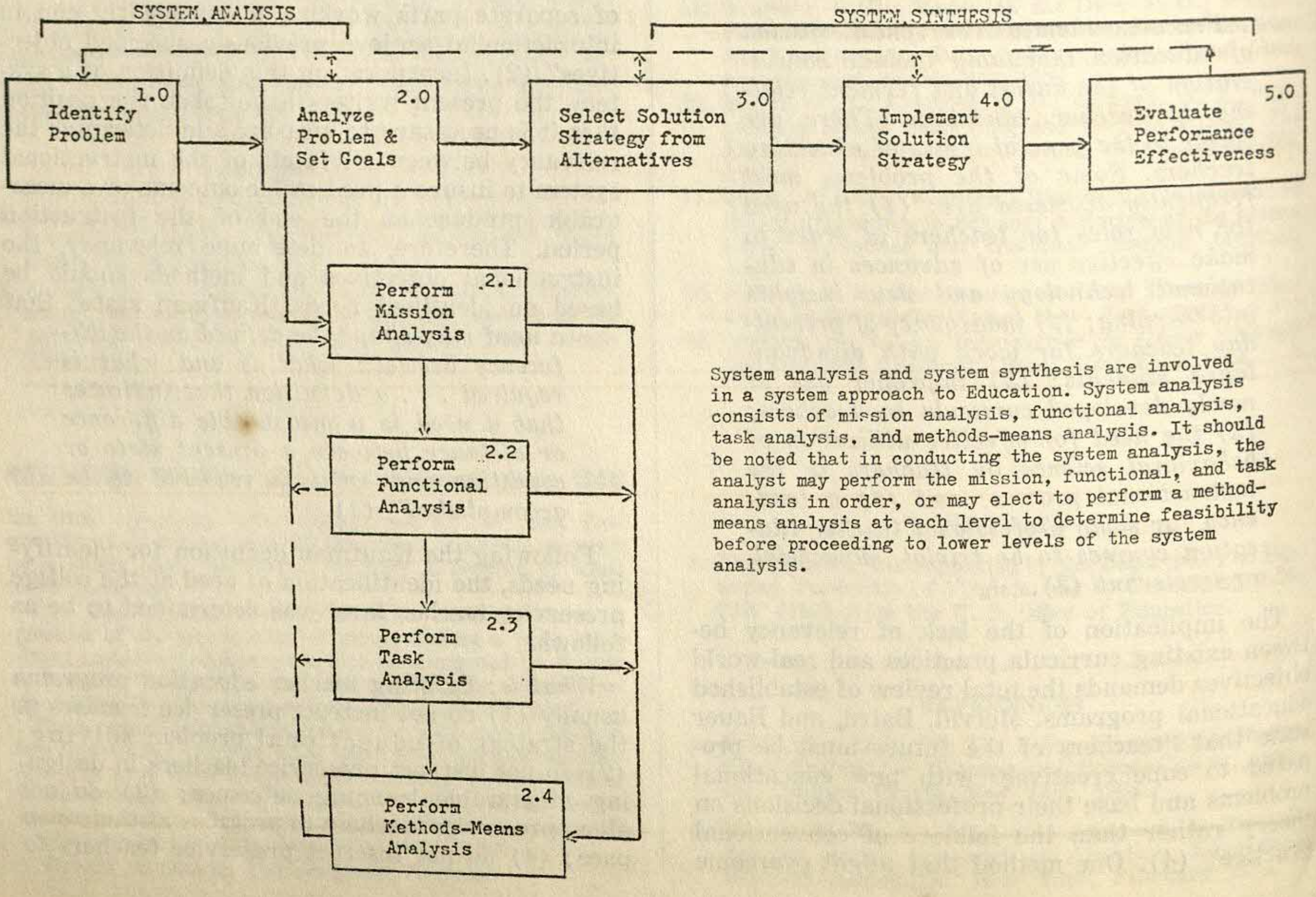
Based on the "what is" and "what is required," the preservice teacher who graduates under existing curriculum programs usually will lack the technological ability necessary to construct and manage an instructional environment conducive to measurable learning. Changing the existing paradigm to incorporate the system approach, then, would facilitate the possibility of determining the interrelating functions and characteristics of a *complete* instructional design and act as a viable model incorporating relevant change.

Description of the Program

The system approach to instructional technology was introduced into the fall quarter curriculum, 1968, and continued through the winter and spring quarters. The basic system course was designed according to Kaufman's paradigm, shown in Figure 1.

Preservice teachers registered for two required Education courses, 430 and 450. One section of each was randomly selected to serve as the experimental section. All preservice teachers in the selected class were designated as part of the experimental group and were instructed in the systems approach to education as described below. The remaining students, comprising part of the control group, were instructed in the traditional lecture approach to (1) preparing lesson plans, (2) writing teacher-made tests, (3) formulating objectives, and (4) selecting audio-visual equipment. The course was intended to instruct

FIGURE 1. System Analysis and System Synthesis


SYSTEM ANALYSIS
SYSTEM SYNTHESIS
1.0
Identify Problem
2.0
Analyze Problem & Set Goals
3.0
Select Solution Strategy from Alternatives
4.0
Implement Solution Strategy
5.0
Evaluate Performance Effectiveness
2.1
Perform Mission Analysis
2.2
Perform Functional Analysis
2.3
Perform Task Analysis
2.4
Perform Methods-Means Analysis

System analysis and system synthesis are involved in a system approach to Education. System analysis consists of mission analysis, functional analysis, task analysis, and methods-means analysis. It should be noted that in conducting the system analysis, the analyst may perform the mission, functional, and task analysis in order, or may elect to perform a method-means analysis at each level to determine feasibility before proceeding to lower levels of the system analysis.

learners in the development of an instructional design based on identified needs. It included the diagrammatic analysis of inherent problems and the procedure in developing relevant solution strategies, rather than the performance of a sophisticated mathematical analysis of existing educational systems. Consequently, the analysis phase consisted of having the learner perform a mission analysis, functional analysis, task analysis, and methods/means analysis, on a mission objective relating to the preparation of instructional materials, from the learner's major field of study.

A second course demanded the design of a 10-hour individual instructional packet developed from the mission analysis, as part of the solution phase. The synthesis course was conducted as a modified individual program, complete with pretest, objectives, learning activities, and method for evaluation included in the student manual. This enabled the preservice teachers to progress at their own rate *within* the restrictions of a quarter time-designation. Emphasis for the 10-hour packet included (1) pretest—posttest construction; (2) written behavioral objectives; (3) designing learning activities, incorporating different learning styles; and (4) writing a programmed instruction sequence. The design of the course required changing the role of the teacher from the "information source" to an instructional leader whose major responsibility was designing and implementing the course.

Statement of Purpose

The purpose of this study was to determine the effect of training in the system approach upon changing preservice teacher attitudes toward the following instructional design factors: (1) product oriented instruction; (2) criteria measure tests; (3) need assessment; (4) system approach; (5) task analysis; (6) method media analysis; (7) learning style; (8) individual instruction; (9) programmed instruction; (10) behavioral objectives; and (11) lecture. Specifically, the following hypotheses were tested:

Hypothesis 1—*There is no significant difference between experimental and control group attitudes toward selected design factors as measured by the evaluation scale adjectives of the semantic differential.*

Hypothesis 2—*There is no significant difference between experimental and control groups' overall attitude toward the program as measured by: (A) the evaluating adjectives of the semantic differential; (B) the receptivity adjectives of the semantic differential; (C) the stability adjectives of the semantic differential.*

Population and Sample

The population for this study consisted of all students accepted into the teacher education program at Western Montana College. Students who had taken the two courses—Education 430, the analysis phase of the system approach, and education 450, the solution phase—were designated as the experimental population. Students who had not taken both of these courses were designated as the control population. A random sample of each of the two groups was chosen using the

Table 1. Median Test Analysis of Overall Student Attitude Toward the System Analysis Program as Measured by the Semantic Scales Evaluation, Stability, and Receptivity

	Design Factors*													
Scale	1	2	3	4	5	6	7	8	9	10	11	df	X^2	p
Evaluation												1	11.18	.001
Experimental	+	+	+	+	+	+	+	+	+	+	–			
Control	–	+	–	–	–	–	–	–	–	–	–			
Stability												1	3.06	----
Experimental	+	+	+	+	+	+	+	+	–	–	–			
Control	+	+	+	+	+	–	–	–	–	–	–			
Receptivity												1	3.06	----
Experimental	+	+	–	+	+	+	+	–	+	+	–			
Control	+	–	+	–	+	–	+	–	–	–	–			

*In same order as listed under statement of problem.

Yates Table of Random Numbers. Twenty students were selected as the experimental group and twenty-three students as the control group.

Method

A semantic differential was constructed measuring eleven design factors relative to three semantic scales: (1) evaluation, (2) stability, and (3) receptivity. Five bipolar descriptive adjectives were used to measure the students' attitude on evaluation, one adjective for stability, and one for receptivity. The students responded on a 7-point scale ranging from very closely related to a neutral position in relation to one end of the scale.

A posttest, experimental-control group design was utilized for this study. The semantic differential was given to the students at the end of spring quarter, 1969, to determine their attitudes toward the experimental program.

The overall student attitude toward the program was determined by evaluating the specified design factors as measured by the semantic scales of evaluation, receptivity, and stability. These data were then analyzed by the median test utilizing the chi-square statistical technique to determine significant differences between the two samples studied.

The student attitude toward each individual design factor was indicated by his rating of the descriptive adjectives under evaluation. An evaluative score was computed on each concept for each S by averaging the evaluative scales. The direction of change for each concept was indicated by a + or — sign. The sign (+ or —) represents the direction of change from the median, a plus (+) sign indicating the mean judgment of the rating to be more favorable and conversely for the minus (—) sign. The median test for K-independent samples statistical technique was again utilized to analyze the data. Significant differences between the two samples were determined by treating the results of the median test to the chi-square technique with the Yates correction factor for small expected frequencies.

Significant differences were accepted when the level of confidence was at .05 or less. The data concerning the factor "lecture" were reversed because negative ratings were the experimenters' positive ratings.

Results

Table 1 shows a significant difference (p. <001) existed between the two groups' attitudes toward the overall program, as measured by the semantic scale, evaluation. Positive attitudes, though not statistically significant, were reported between the two group attitudes toward the overall program by the stability scale and receptivity scale.

Table 2. Median Test Analysis of Student Attitude Toward Selected Design Factors as Measured by the Semantic Scale Evaluation

	EVALUATION[a]			
Instructional Design Factor	Experimental Median N=20	Control Median N=23	Combined Median N=43	X^2
Product Oriented Instruction	6.3	5.8	5.0	6.4**
Criteria Measure Tests	6.2	5.8	5.8	2.6
Need Assessment	6.1	5.6	5.8	.28
System Approach	6.1	5.6	5.8	.00
Task Analysis	5.9	5.2	5.6	.28
Method/ Media Analysis	5.9	5.2	5.6	6.4**
Learning Styles	5.9	5.2	5.6	4.5*
Individual Instruction	5.8	5.2	5.6	.28
Programmed Instruction	5.8	5.4	5.4	6.4**
Behavioral Objective	5.7	5.2	5.4	4.5*
Lecture	2.6	3.6	3.0	4.5*

**p .02
*p .05
[a] df =1

As indicated in Table 2, significant differences were found between the experimental and control group evaluative ratings for (1) product oriented instruction, (2) method/media analysis, (3) learning styles, (4) individual instruction, (5) behavioral objectives, and (6) lecture, after completing the instructional program. Both groups rated all the factors except lecture favorable. Lecture was rated as an unfavorable instructional factor.

The results of the findings indicated:

(A) *Evidence to reject hypothesis 1. There existed a significant difference between the experimental and control groups' attitudes toward six of the selected designated factors as measured by the semantic scale evaluation.*

(B) *Evidence to reject hypothesis 2A. There did exist a significant difference (.001) between the experimental and control groups' attitudes as measured by the evaluation semantic scale.*

(C) *No evidence to reject hypothesis 2B and 2C.*

Discussion

Both the experimental and control groups tended to rate the experimental program favorably. The experimental group rated the majority of the design factors significantly higher than did the control group. Teaching preservice teachers the analysis and solution phases of the system approach to education can change preservice teachers' attitudes toward various design factors. An introduction to the change of the education paradigm must begin with a change in teacher attitude.

We may conclude from this study that system approach concepts can affect preservice teacher attitudes from a teacher-oriented to a learner-oriented instructional environment.

Teacher training institutions should study the possibilities of utilizing the system approach in training teachers to overcome the characteristics of the education courses that merit such criticism as trivial, superficial, or irrelevant—and produce a teacher that can make instructional decision based on theory, not folklore.

REFERENCES

1. Kaufman, R. A., "A System Approach to Education," *AV Communication Review*, 16:(no. 4), Winter 1968.
2. Kaufman, R. A.; Corrigan, R. E.; Corrigan, B. O.; Goodwin, D. L., "Steps and Tools of System Analysis as Applied to Education"; "Steps and Tools of the System Synthesis Process in Education"; "Task Analysis in Education"; "Methods-means Analysis in Education"; "Mission Analysis in Education"; "Functional Analysis in Education"; "An Interim Generic Problem Solving Model"; "An Exercise in the Analysis of Planned Change in Education." (eight pamphlets prepared for Operation PEP), San Mateo County Department of Education, San Mateo, California, 1967.
3. Lange, Carl J., "Developing Programs for Teachers," *Preparing Educators to Meet Emerging Needs* (Part I), Designing Education for the Future, An Eight State Project, Denver, Colorado, p. 112.
4. Merrill, David; Baird, Hugh; Bauer, Edith, *Teacher Education 1984*, Brigham Young University, Provo, Utah, 1966.

New Books from the Publishers

Elementary

Almy, Millie and others. *Logical Thinking in Second Grade* (New York: Teachers College Press, 1970), xiii + 216 pp. Paper $2.95.

General

Bayliss, Sylvia J. and others (eds.). *Career Opportunities: Community Service and Related Specialists* (Chicago: J. G. Ferguson Publishing Company, 1970), xvii + 292 pp. Cloth $11.95.

Blishen, Edward (ed.). *Encyclopedia of Education* (New York: Philosophical Library, 1970), xi + 882 pp. Cloth $20.

Elwood, Bryan C. *Student Transportation: Comparing Alternative Methods of Providing the Service* (Toronto: The Ontario Institute for Studies in Education, 1970), vii + 24 pp. Paper $2.

Grusky, Oscar and George A. Miller (eds.). *The Sociology of Organizations: Basic Studies* (New York: The Macmillan Company, 1970), x + 592 pp. Cloth $11.95.

Johnson, Marjorie Seddon and Roy A. Kress (eds.). *Thought–Language–Reading: Proceedings of the 1969 Annual Reading Institute at Temple University* (Philadelphia: The Reading Clinic, Temple University, 1970), 108 pp. Paper $4.50.

Kinsinger, Robert E. (ed.). *Career Opportunities: Health Technicians* (Chicago: J. G. Ferguson Publishing Company, 1970), xviii + 386 pp. Cloth $11.95.

Loree, M. Ray. *Psychology of Education* (second edition) (New York: The Ronald Press Company, 1970), viii + 666 pp. Cloth $9.

McAshan, H. H. *Writing Behavorial Objectives: A New Approach* (New York: Harper and Row Publishers, 1970), xi + 116 pp. Paper $2.50.

Marx, Melvin H. (ed.). *Learning: Theories* (New York: The Macmillan Company, 1970), x + 340 pp. Cloth $9.95.

Pintner, Rudolf and others. *Educational Psychology* (sixth edition) (New York: Barnes and Noble, Inc., 1970), xvi + 308 pp. Paper $1.75.

Smith, Carl B. *Treating Reading Difficulties: The Role of the Principal Teacher, Specialist, Administrator* (Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1970), 132 pp. Paper $1.50.

Wilson, J. Donald, Robert M. Stamp, and Louis-Philippe Audet (eds.). *Canadian Education: A History* (Scarborough, Ontario: Prentice-Hall of Canada Ltd., 1970), xiv + 528 pp. Cloth $9.95.

Teaching

Cuban, Larry. *To Make a Difference: Teaching in the Inner City* (New York: The Free Press, 1970), xxii + 261 pp. Cloth $6.95. Paper $3.50.

Greenberg, Selma. *Selected Studies of Classroom Teaching: A Comparative Analysis* (Scranton, PA.: International Textbook Company, 1970), xi + 160 pp. Cloth 6.50.

Kaye, Barrington. *Participation in Learning: A Progress Report on Some Experiments in the Training of Teachers* (London: George Allen and Unwin Ltd., 1970), 239 pp. Cloth 40 shillings.

Phillips, Gerald M., Robert E. Dunham, Robert Brubaker, and David Butt. *The Development of Oral Communication in the Classroom* (Indianapolis: The Bobbs-Merrill Company, Inc., 1970), xi + 211 pp. Cloth $6.50.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

A Study of Children's Drawings of Their Classrooms[1]

R. S. ROGERS and E. N. WRIGHT[2]

ABSTRACT

This report summarizes a large scale study of children's drawings of their classrooms conducted by the Toronto Board of Education. The drawings were collected under standardized conditions from junior kindergarten to grade 4. A taxonomic approach to the drawing content through an objective coding framework is described and summaries of *ad hoc* examinations of the data are given along with experimentation involving the actual drawings. The discussion suggests that the findings of the study are parallel to those in the literature on children's drawings as a whole, namely, that measures of skill of execution ("drawing intelligence") are more stable and reliable than measures of inferred mood, attitude, or personality.

A RECENT article in this Journal by Gregersen and Travers (5) drew attention to the fact that "Relatively few studies have explored the perceptions which elementary school pupils have of their classrooms and their teachers." The article, however, confined itself to a consideration of the child's perception of his *teacher* as inferred from drawings by two judges on the basis of a preestablished coding framework. The present report seeks to provide information on the other aspect of school perception to which attention was drawn by these authors, namely the child's view of his or her *classroom.*

Despite a careful search of the voluminous literature on children's drawings, the senior author (14) and two of his predecessors in the Research Department of the Toronto Board of Education (2, 9) have failed to find reference to any previous research on children's drawings of their classrooms. Consequently, an exploratory as opposed to an hypothetico-deductive approach was adopted. The coding of the drawings, analyses of the differential occurrence of these codes in various groups, and other experimentation were evolved based on experience (see 17 for example).

Method

Each teacher gave the following instructions to the class:

> *I am going to ask you to do something very special for me. I want you to make a picture of our room. Look all around the room and then draw a picture of our room on the paper.*

The teachers were instructed not to "coach" but to give encouragement only by their attitude. The children were provided with a large piece of newsprint and crayons in eight specified colors with which to make the drawings. When a child had finished the teacher said: "Tell me about the picture you have made." The teacher recorded verbatim the children's responses to what they had drawn, recording the words close to the content that the children pointed out.

Subjects. The Ss were 8,695 pupils who entered senior kindergarten in Toronto in 1961. This population was part of a longitudinal study of which the collection of the drawings was only one part. The number of drawings collected from this population follows:

Time of Collection	*Grade Position Given Normal Promotion*	*Drawings (to nearest 100)*
March 1961	Junior Kindergarten	1500
April 1961	Junior Kindergarten	1400
March 1962	Senior Kindergarten	8400
May 1962	Senior Kindergarten	8300
April 1963	Grade 1	6800
February 1964	Grade 2	6400
March 1965	Grade 3	5800
April 1966	Grade 4 (only a sub-sample tested)	1600

The drawings collected in junior kindergarten were from those in the population who were in junior kindergarten prior to entering senior kindergarten (approximately one-sixth of the students attended junior kindergarten). The drawings collected in grade 4 were from only a sample of the students. A total of some 40,000 drawings was collected and later transferred to color microfilm for permanent storage.

Scoring Procedure. The categories for coding the drawing content were established by a team of coders averaging six to seven in number. The coders had no personal knowledge of the children involved and came from a variety of national and academic backgrounds. "Bias" was therefore felt to be at a minimum.

A record was made of the frequency with which various kinds of content appeared. As each code was developed, it was discussed among the coders to insure that its wording was not perceived ambiguously and that it did not replicate another code. Reliability trials used 82.5 percent as a criterion, i.e., there had to be complete agreement among five raters on coding 100 out of every 120 sample drawings or the code and instructions were redefined.

As the various aspects of drawing content so abstracted were examined, it became clear that an objective analytic framework must revolve around three basic aspects of content:

Objects: *what objects, if any, are present;*
Persons: *what persons, if any, are present;*
Space: *how the objects and/or persons are presented and disposed.*

The final coding framework consisted of the following specific aspects of content:

- (i) *Space*
  - Framing or Enclosure
  - Partitioning
  - Physical Boundaries
  - Amount of Space Used
  - Scope or Subject Matter
  - Shading
  - Compression or Expansion
  - Viewpoint
- (ii) *People*
  - (a) Persons
  - Appearance of Boy(s)
  - Appearance of Girl(s)
  - Appearance of Unspecified Children
  - Sex of Unspecified Children
  - Size
  - Activity and Interaction
  - Labeling
  - (b) Drawing the Person
  - Head
  - Facial Features (Stereotyping)
  - Neck
  - Arms
  - Fingers, hands
  - Legs
  - Feet (Position, Proportion)
  - Trunk (Stereotyping)
  - Body Representation
  - View of Person
  - Clothing
  - Transparency
  - Detail
  - Use of Color
- (iii) *Objects*
  - (a) Objects
  - Transparency
  - Fold Out
  - Behind, In Front Of
  - On-Under
  - Perspective
  - Predominant Color
  - Time by Tense of Label(s)
  - Labeling to Indicate Use
  - Time by Pictorial Situation
  - Classroom Atmosphere
  - (b) Classroom Constants
  - Inclusion
  - Shading
  - Use of Color
  - Grouping of Types of Constants

Within each of the three content areas, there were some fifteen to twenty different subareas as have been listed and within each of these subareas, two to ten individual codes. For example:

Content Area	*Subarea*	*Individual Codes*
Space	Joined Frame Lines	no frame lines present only one frame line present more than one frame line present but not joined more than one frame line present and joined.

Each drawing was classified as belonging in one but not more than one individual code within each subarea. Complete details of the codes were available to the coders at all times through a mimeographed coding manual (8). As may be seen from the example above, considerable efforts were made to define and objectify content. Three tests of inter-coder reliability were made at different time points during the coding (14). These tests indicated that two coders working independently agreed as to code at least nine times in ten.

Summary of Experiments

Because of limitations of time, only a sample of some 3,500 drawings was completely coded. There was thus available for study both a pool of data based on this sample and the 40,000 original drawings. *Ad hoc* investigations were made using these two sources; these are now summarized.

Work conducted in the early days of the project (2) on the "retest reliability" of drawings, that is the extent to which there are continuities in the drawings of individual children, may be seen also as a measure of the extent to which each child maintains a constant vision of his or her classroom. In brief, this work involved examining the extent to which the content observed in one pupil's drawing made in kindergarten was the same as that in another drawing made by the same child a month later. In total, four hundred such pairs of drawings were examined, two hundred made in junior kindergarten and two hundred made in senior kindergarten. Where there were human figures in each of a pair of drawings they were given Goodenough-Harris point scores (6), the average correlation across pairs was +0.85 for the junior kindergarten drawings and +0.91 for the senior kindergarten drawings. For the content coded according to other categories, however, the resemblance between a pair of drawings was considerably less and only an average of 56 percent of codes were identical across pairs. Measures of techniques of execution, e.g., using color to differentiate parts of an object were more consistent than measures of specific content such as "presence of blackboards."

Study 1. In one *ad hoc* study, consideration was given to the effect that the age of the "artist" had on the drawing content as revealed by the coding framework. A random sample of one hundred children was drawn from the population of children who had completed a drawing at every stage from junior kindergarten to grade 4. It was thus possible to study changes in the drawing content of the same one hundred children across time.

Many of the codes showed considerable time-linked changes. As these time trends were considered, it became clear that three different kinds of changes seemed to dominate the results. These were:

> Realism—*Objects looked more and more as they would in a photograph.*
> Sophistication—*an increasing skill at differentiating objects and their interrelationships was displayed.*
> Conformity of Content—*an increasing tendency to conform to the letter of the instructions, (i.e. to draw the classroom and nothing else) appeared.*

The changes in content which were felt to support the above synthesis are reported elsewhere (12).

Study 2. The second major exploration of the data pool (13) considered whether drawings made in senior kindergarten were different if the child had attended junior kindergarten. A "matched pair" design was used for each child who attended junior kindergarten, a child who did not go to junior kindergarten was selected on the basis of the following matching factors: sex, age, language(s), father's and mother's education, and father's occupation. Two separate data analyses were made, one in which 558 matched pairs were obtained where the children who did not go to junior kindergarten nonetheless lived in areas of Toronto where junior kindergarten was available: in the other set of 556 matched pairs the children who did not attend junior kindergarten who lived in areas of Toronto where junior kindergarten was not available. The data from both senior kindergarten drawing collection times were considered. A test of proportions (13) was used to identify codes that had differing frequencies of occurrence between "attenders" and "non-attenders" of junior kindergarten. Although a number of minor differences were detected in both comparisons, these were less marked for the second set of senior kindergarten drawings than for the first. This was seen as evidence that the effects of junior kindergarten were diminishing over the month interval between the two drawing collections.

Study 3. It was observed that in some of the kindergarten drawings the children had included number symbols. Another study was, therefore, undertaken to determine if the presence of these numbers in the drawings, before their formal introduction into the curriculum, would serve as an indicator of a child's mathematical achievement in grade 1 (4).

To answer this question, pairs of students were selected and matched on a number of criteria (sex, age, school experience, and socioeconomic status). This procedure resulted in two groups of forty-three students with the only apparent difference between them being that one group had included numbers in their kindergarten drawings while the other group had not.

All Ss were given a set of tests in grade 1 (a year after completing the drawings) to determine their level of understanding of numeric concepts. One test, the Metropolitan Achievement Test, is commercially available and is traditionally used to measure arithmetic achievement. The other test, a more novel one, was developed by Dodwell (3), a Canadian experimental psychologist, on the basis of work done by Piaget (10). In essence, this test attempts to measure the degree to which the child is able to understand the underlying principles of mathematics, quantity, and number. It was found that there were no meaningful differences between the groups on the basis of the tests.

On the basis of these results, it was concluded that the spontaneous production of number symbols in the classroom drawings of kindergarten children could not serve as an indicator of their number concept understanding in grade 1.

Study 4. Another area of work with the drawings (14) resulted from the feelings expressed by many of the coders and other workers, who were exposed to the drawings, that some of the graphic output was indicative of various kinds of maladjustment in those children. This observation coupled with the affinity of the drawings to those used in such diagnostic tools as the "Draw-a-Person" test of Machover (7) or the "House-Tree-Person" test of Buck (1) led to the decision to evaluate the usefulness of the drawings as indicators of later referral to Toronto's Child Adjustment Services (CAS).

In brief, pairs of drawings made in kindergarten were presented in turn to five judges with levels of prior experience with children's drawings varying from considerable to effectively nil. In each pair of drawings, one drawing was from a child who had never been referred to CAS while the other came from a child who had been referred under one of the following broad categories:

(a) Developmentally/Functionally Subnormal, e.g., poor school progress, language problem, mentally retarded, marked immaturity;

(b) Projected Personality Problem, e.g., lying, stealing, fighting;

(c) Introjected Personality Problem, e.g., nail biting, timidity, daydreaming, bizarre behavior;

(d) Poor Social Relationships, e.g., inability to get on with others, overly aggressive, withdrawing behavior;

(e) Environmental Disadvantage, e.g., poor home conditions, poor family relationships, attendance problem;

(f) Intellectually above Normal, e.g., referred for psychometric test as felt by teacher to be superior.

The judges were told into which of the six broad categories one of a pair of drawings fell; their task was to say which one (e.g., "either drawing #7799 or drawing #7793 came from a child referred under category a; which one was it?").

Summarizing the results of all judges over all parts of drawings, there were 479 correct judgments and 370 incorrect judgments. This yielded a X^2 of 13.7 ($p < .001$, $df = 1$). Hence, a low (13 percent above chance) but statistically reliable degree of discrimination was obtained. It was also noted that:

(1) The ability to discriminate did not seem to be appreciably affected by the judges' presumed familiarity with children's art;

(2) Average discrimination scores were probably higher in the referral categories concerned with ability than in categories concerned with personality or environment.

The first and major finding, indicating that it was possible to detect some of those who were later referred to psychological services by examination of their drawings is congruent with the research of Westman and others (16). These workers used human figure drawings as their index. The senior author has reported elsewhere (15) that referrals, with the exception of a few who are sent for being "bright," are generally poor performers on academic tests. As discrimination of referrals was better for categories a and f, which dealt with ability, it would seem reasonable to hypothesize that where the judges could discriminate, the cues were related to skill in executing the drawing.

Conclusions

The possibility of establishing relationships among various characteristics of the drawings and other educational variables concerned with school achievement was approached from two directions:

(a) The examination of the drawings of groups of students known to differ in educational experience;

(b) The attempt to predict academic differences from drawing content.

The first direction proved more fruitful: the drawings showed distinct changes as the children got older, and less dramatic differences were noted between the drawings of children who attended junior kindergarten and those who did not. By way of contrast, the second approach indicated that the appearance of numerals in drawings was no guide to later mathematical achievement and that only the most marginal success would be likely to accompany an attempt to use the drawings as predictors of later referral to CAS.

It was also noted that when the drawings of the same child made at a 1-month interval were considered, the continuity was high as far as aspects of overall skill of execution were concerned but much lower for items of specific content.

Discussion

The research reported here summarizes some 8 years of investigational effort. During that time, four researchers have made separate considerations of the literature (2, 9, 11, 14). The senior author, who worked with the project in its final year, was therefore in a very favorable position to attempt a synthesis of the findings of the project with the literature on other research with children's drawings. The outcome of this attempt was the conclusion that research on graphic output can be classified as either seeking to measure skill of execution or attempting to infer "meaning" (where by meaning we understand such things as inferences about "projections of the artist's personality" or what the artist "sees"). The well researched Goodenough-Harris "Draw-a-Man" test leaves little doubt that skill of execution is a good predictor of the results of more formal measures of intelligence. The same cannot be said for at least the vast majority of attempts to validate the inference of mood, attitude, personality, or temperament from drawings.

Within this synthetic framework, the data reported here seem to make good sense. Measures of skill of execution were found to be more constant across time than measures of specific items of content. The drawings were not found to be very efficient diagnostic tools for children referred for "personality" problems. Most of the changes noted in drawing content across time were classifiable as aspects of increasing skill. The appearance of numbers in a child's drawings was not found to be an index of his or her mathematical achievement.

As a result of these considerations, the author, in presenting the outcome of the research to teachers (14), has emphasized that drawings are best seen as an informal focal point from which communication can be started, rather than as cryptic clues to the inner world of the child awaiting some "magic" interpretative code for their full understanding.

FOOTNOTES

1. This report deals with a small portion of data collected as part of a longitudinal study of achievement. This study examines factors in the home, socioeconomic status, mother tongue, parental expectations, and attendance at junior kindergarten as they relate to school achievement. Information and reports are available from the Research Department of The Board of Education for the City of Toronto.

2. Appreciation is expressed to Dr. A. R. MacKinnon, Pramuan Dickinson, Judith Palmer, S. S. Reider, Sheila Pennington, and James Feeney whose work on the project made this summary report possible.

REFERENCES

1. Buck, J. N., "The H.T.P. Technique: A Quantitative and Qualitative Screening Manual," *Clinical Psychological Monograph*, 5:1-120, 1948.

2. Dickinson, Pramuan, "A Study of Drawings Made by Children of Junior Kindergarten, Senior Kindergarten, and First Grade of Toronto Public Schools," unpublished EdD thesis, University of Toronto, Toronto, Canada, 1965.

3. Dodwell, P. C., "Children's Understanding of Number Concepts: Characteristics of an Individual and of a Group Test," *Canadian Journal of Psychology*, 15: (no. 1)29–36, 1961.

4. Feeney, J. D.; Wright, E. N.; Adler, Marilynne, *Children's Concept of Number: The Spontaneous Production of Number Symbols in Their Drawings*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1968.

5. Gregersen, Gayle F.; Travers, R. M. W., "A Study of the Child's Concept of the Teacher," *The Journal of Educational Research*, 61:(no. 7)324–327, 1968.

6. Harris, D. B., *Children's Drawings as Measures of Intellectual Maturity*, Harcourt, Brace, and World, New York, 1963.

7. Machover, Karen, "Human Figure Drawings of Children," *Journal of Projective Techniques*, 17:(no. 1) 85–91, 1953.

8. Pennington, Sheila, *Draw-A-Classroom "Test" Manual and Scoring Categories*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1966.

9. Pennington, Sheila, *Children's Drawings: A Selected Review of Literature*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1967.

10. Piaget, J., *The Child's Conception of Number*, Translated by C. Gattegno and F. M. Hodgson, Routledge and Kegan Paul, London, England, 1952.

11. Reider, S. S., *Artwork: A Brief Review*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1967.

12. Rogers, R. S., *Changes With Time in the Content of Children's Drawings—A Longitudinal Study With the "Draw-A-Classroom Test,"* The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1968.

13. Rogers, R. S., *The Effect of Having Previously Attended Junior Kindergarten on "Draw-A-Classroom" Test Scores Obtained in Senior Kindergarten*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1968.

14. Rogers, R. S., *A Consolidated Report on the "Draw-A-Classroom" Test—A Study of the Drawing Behaviour of Children in Toronto Public Schools*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1969.

15. Rogers, R. S., *Referrals to Child Adjustment Services: A Longitudinal Study Using Data From the Study of Achievement*, The Board of Education for the City of Toronto, Research Department, Toronto, Canada, 1969.

16. Westman, J. C.; Ferguson, Barbara B.; Wolman, R. N., "School Career Adjustment Patterns of Children Using Mental Health Services," *American Journal of Orthopsychiatry*, 38:(no. 4)659–665, 1968.

17. Wright, E. N., "The Development of the Draw-A-Classroom Test: A Case Study," *Proceedings—Third Canadian Conference on Educational Research*, June, 1964, pp. 203–208.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 8, April 1971)

An Empirical Comparison of Selected Enrollment Projection Formulas

WILLIAM J. WEBSTER
Director, System-Wide Evaluation, Dallas Independent School District

ABSTRACT

The purpose of this study was to determine if the accuracy of enrollment projections is systematically related to the use of different selected ratio and regression projection formulas, to differences in community growth characteristics, and to differences in the educational level for which projections are made. Seven projection formulas, including those currently most used in education, were compared across communities with different growth characteristics, for different grade levels, for a sample of forty Michigan school districts. Results suggest that there are systematic relationships between accuracy of enrollment projection and those variables cited above.

IN RECENT years the need for greater precision in the projection of future school enrollment has become increasingly apparent. However, the interaction of a number of diverse demographic variables, by no means uniform from school to school, has contributed to making the desired precision difficult to attain. Also contributing to the current lack of precision in educational attendance projection has been a general lack of imagination among most practitioners with regard to developing new projection techniques that more closely meet the problems of public school attendance projection than those currently in use. Specifically, little experimentation has been done in a public school setting with any projection techniques other than ratio methods, that is, methods that employ the ratio or proportion of a predictor to a criterion at some time in the past as the basis for projections (1, 9, 14).

The most sophisticated research on the efficacy of non-ratio projection methods for predicting future population has been carried out by the United States Census Bureau. Specifically, Census Bureau experimentation with a number of types of mathematical functions suggested that, despite the relative accuracy of projections yielded by logistic functions and linear least-squares projection equations, such functions were generally unsatisfactory for projecting future population (6). It was felt that mathematical functions could not depict the year to year variations which result from fluctuations in migration since such functions present only a theoretical pattern of growth (6, 8). For this reason, the Census Bureau utilizes methods which involve the use of vital statistics (birth and death statistics) in conjunction with previous census enumerations to project future population (11). It is significant to note, however, that the United States National Center for Educational Statistics utilizes mathematical functions in projecting future enrollment (12).

The failure of most educators to use various mathematical functions to project school enrollment is perhaps due to an apparent failure to appreciate the differences in problems faced by the Census Bureau in projecting future population and that faced by the school principal in projecting school enrollment. Federal demographers, working with large populations and geographic areas about which a great deal of reliable data is available, have much more information at hand than the educator who must work with small populations and geographic areas about which little reliable data is available. By utilizing only ratio techniques, educators forfeit many of the advantages of the Census Bureau's techniques while retaining the majority of the disadvantages. Where complete and reliable vital statistics and migration data are not available, it would seem that certain varieties of mathematical functions, particularly a linear least-squares equation, would be appropriate. This view has been supported within the context of projecting both university (2) and public school (13) enrollment.

In addition to the aforementioned concerns with regard to projection methodology, it has been suggested that different projection methods might be applied to districts with different demographic characteristics in order to obtain more accurate projections (5). The present study was designed to investigate the above mentioned concern as well as a number of related ones. In terms of accuracy of projections, eight research questions were of interest. They were, are there systematic effects due to:

1. the use of selected different projection formulas?
2. differences in community growth characteristics?
3. differences in educational level?
4. the combination of particular projection formulas with particular community growth characteristics?
5. the combination of particular projection formulas with particular educational levels?
6. the combination of particular educational levels with particular community growth characteristics?
7. the combination of particular projection formulas with particular community growth characteristics with particular educational levels?
8. the use of linear regression as opposed to the use of traditional ratio methods?

Methodology

Seven different ratio and regression methods were employed in this study. The distinction between ratio and regression methods is made on the basis of the measure of relationship used in the computation of projected enrollment. Ratio methods employ the ratio or proportion of a predictor to a criterion in the past, while regression methods depend upon coefficients of correlation and multiple correlation, statistics which purport to measure degree of association. The three ratio methods used were those identified by Hauser and Kitagawa (5), MacConnell (9), and others (1, 14) as those most used by educators. Regression analyses, utilizing the same predictors as each of the three ratio methods, were also employed. In addition, a regression approach using time as a predictor was used.

The reader who is unfamiliar with the problems involved in projecting local school enrollment may wonder why separate and distinct regression equations were used instead of one of a variety of procedures that take a number of predictor variables and select from them the "best" regression equation (4). There were two major reasons why such an elimination procedure was not used. First, previous research by the author suggested that the efficacy of a given set of predictor variables for projecting enrollment could not be predicted from the correlation of each of those variables with the criterion, either individually or as a composite (14). The reason for this phenomenon is that in projecting enrollment one must extrapolate the projections beyond the domain of the data on which those projections are based. It is therefore essential that the variables used in a projection equation not only correlate highly with the criterion, but that they also exhibit a robustness in being extrapolated. Although a correlation of approximately .70 between a predictor and a criterion would seem to be a necessary attribute for a variable's use in projection, it is certainly not a sufficient one. A given variable or set of variables must be empirically tested for a given population before any confidence can be validly placed in their use.

Second, elimination procedures generally require more degrees of freedom than are usually available when projecting local school enrollment. The present study used only ten data points. Since high multiple correlation coefficients appear as artifacts of the computational process as the number of predictor variables approach the number of data points, an elimination procedure was not seen as meaningful. Adding data points was not a viable alternative since unusual fluctuations in population occurred as a result of the Second World War, and therefore, by adding data points the amount of error in the projections would have been increased.

Table 1 presents the seven projection methods employed in this study in terms of whether each is a ratio or regression method, the predictor variable or variables used in each, the relationship on which each projection was based, and the source from which the reader can obtain a detailed specification of the methodology involved. In each case thirteen separate equations were used to project enrollment for each school district by each formula. That is, one formula was developed for each methodology for projecting enrollment in each grade. The grade projections were subsequently combined to obtain projections of elementary (K—8) and secondary (9—12) enrollment.

The population utilized in this study consisted of every public school district in the state of Michigan that had grades K-12 and was largely coterminous with a city of at least 10,000 population. The population was divided into five strata on the basis of the following community growth characteristics:

1. *the percent of increase in general population, during the decade 1950–1960, of the urban place with which the majority of each school district was geographically coterminous;*
2. *the percent of increase in the number of households during the decade 1950–1960, of the urban place with which the majority of each school district was geographically coterminous;*
3. *the population density increment of*

the urban place with which the majority of each school district was geographically coterminous. The population density increment was computed by subtracting the 1960 population density of the county in which the district was located from the 1960 population density of the city with which the majority of the district was geographically coterminous.

A sample of eight school districts was randomly drawn from within each of the five strata, yielding a total sample of forty school districts. Table 2 outlines the decision rules used in classifying districts into strata.

The seven different projection methods described in Table 1 were used to develop projections of K-8 and 9-12 enrollment for the year 1965. These projections were based on data that were available for the years 1950 through 1960.

Table 1.—Projection Methods Employed

Method	Type	Predictor(s)	Relationship and Source
1. Cohort-Survival	Ratio	1. Births by city. 2. Past district enrollment.	Past ratios of birth to kindergarten and grade to subsequent grade enrollment (1, 3, 14).
2. Transition Analysis	Regression	1. Births by city. 2. Past district enrollment.	Past tendencies of students to advance from birth to kindergarten and from one grade to the next (2, 13, 14)
3. Pupil-Population I	Ratio	1. City school-age population.	Past ratio of public school enrollment to city school-age population (5, 14).
4. Pupil-Population II	Regression	1. City school-age population.	Past relationship between public school enrollment and city school-age population (14).
5. Pupil-Home Unit	Ratio	1. City occupied housing units.	Past ratio of public school enrollment to city occupied housing units (5, 14).
6. Housing Analysis	Regression	1. City occupied housing units.	Past relationship between public school enrollment and city occupied housing units (14).
7. Time Analysis *	Regression	1. Time	Past relationship between public school enrollment and time (2, 14).

* No ratio analog to this method employed.

Table 2.—Decision Rules Used in Classifying Districts into Strata

Stratum	% Increase General Population 1950-1960	% Increase Households 1950-1960	Population Density Increment
A	>100	>100	negative
B	>100	>100	positive
	≥ 50 but <100	> 100	positive or negative
	>100	≥ 50 but <100	positive or negative
	≥ 50 but< 100	≥ 50 but <100	positive or negative
	≥ 50 but <100	< 50 but ≥ 0	negative
	< 50 but ≥ 0	≥ 50 but< 100	negative
C	≥ 50 but <100	< 50 but ≥ 0	positive
	< 50 but ≥ 0	≥ 50 but <100	positive
	≥ 10 but < 50	≥ 10 but< 50	negative
D	≥ 10 but < 50	≥ 10 but< 50	positive
	≥ 0 but < 10	≥ 0 but< 10	positive or negative
	≥ 0 but < 10	≥ 10 but< 50	positive or negative
	≥10 but < 50	≥ 0 but< 10	positive or negative
E	negative	negative	positive or negative
	negative	positive	positive or negative
	positive	negative	positive or negative

As a result it was possible to use actual enrollment in the year 1965 as the criterion against which to judge the accuracy of each of the projection methods.

Each of the forty districts that comprised the sample was characterized by fourteen independent projections of different aspects of its enrollment. These fourteen projections included estimates of elementary and secondary enrollment for the year 1965 for each of the seven projection methodologies used. Since it was desired to compare the various projections across districts, a meaningful standard of comparison was needed. A simple measure of the percentage of error present in each projection, termed the W coefficient by the author, was chosen as the standard of comparison. Symbolically, for any given projection:

$$W = \frac{A_{1965} - E_{1965}}{A_{1965}}$$

where:

A_{1965} = Actual elementary or secondary enrollment for a given school district in 1965.

E_{1965} = Estimated elementary or secondary enrollment for a given school district in 1965.

As an examination of the formula for obtaining the W coefficient shows, the smaller the W coefficient the better the estimate. Since the W coefficients can validly be assumed to have been independently drawn from normal treatment populations with equal variances, and to have had independent error components across all pairs of observations, a 7 × 5 × 2 fixed effects analysis of variance design, with main factors crossed, was employed to explore the first seven research questions listed at the outset of this paper. A separate analysis was used to investigate the relative merits of ratio versus regression projection methodology.

Assumptions Inherent in Projection Methods Utilized

The major assumption that was made for all projection methods utilized in this study was that demographic conditions remained stable, or changed at some regular rate, from the period from which data were drawn to the period to which projections were extrapolated. One additional assumption was made for every regression method used. That is, the line of best fit which specified the relationship between predictor and criterion variables was a straight line. This was seen as a reasonable assumption since most functions specifying data of this type can best be approximated over a short period of time or small number of data points by a straight line. Since the regression methods utilized in this study were used in strictly an algebraic manner, with no confidence intervals being established, no additional assumptions were necessary. Additional

Table 3.—Summary of Analysis of Variance for Projection Methodology, Stratum, and Educational Level Using Accuracy of Projection as the Independent Variable

Source	df	MS	f
Projection Methodology	6	.057994	18.54 **
Stratum	4	.055556	17.76 **
Educational Level	1	.041043	13.12 **
Projection Methodology x Stratum	24	.010042	3.21 **
Projection Methodology x Educational Level	6	.007707	2.46 *
Stratum x Educational Level	4	.019101	6.11 **
Projection Methodology x Stratum x Educational Level	24	.014858	4.75 **
Within	490	.003128	
Total	559		

* Significant at the .05 level.

** Significant at the .01 level.

assumptions needed for the purpose of computing confidence intervals are discussed in depth by Cochrane and Orcutt (3) and Johnston (7).

Results

The results of the three-way analysis of variance for projection methodology, stratum, and educational level, using W coefficients as the criterion of accuracy, are reported in Table 3.

The first seven research questions posed at the outset of this study can be tentatively answered in the affirmative ($\alpha \leqq .05$). In order to investigate the relative accuracy of ratio versus regression methods, the eighth question, the following regression model was employed.

$$Y = a_0U + a_1X_1 + a_2X_2 + E_1$$

where:

Y = the W coefficient

U = the unit vector

X_1 = "1" if projection methodology was a ratio method (Cohort-Survival, Pupil-Population I, Pupil-Home Unit), "0" otherwise.

X_2 = "1" if projection methodology was a regression method (Transition Analysis, Pupil Population II, Housing Analysis), "0" otherwise.

E_1 = the error vector $(Y\text{-}\hat{Y})$

a_n = least-squares weights, $n = 0, 1, 2$

The restricted model was:

$$Y = a_0U + E_2$$

where Y and U are the same as above and

E_2 = the error vector, $(Y\text{-}\hat{Y})$

a_n = least-square weight, $n = 0$

The results of this regression analysis suggest that regression methodology provided more accurate overall projections than did ratio methodology ($\alpha \leqq .01$). The reader should note that Time Analysis was not included because it had no ratio analog.

In order to investigate some of the relationships that contributed to the statistically significant main effects, Scheffé *post hoc* analyses (10) were computed for projection methodology and stratum classification. Table 4 presents the results of the Scheffé test for main effects due to projection methodology. The formulas are arranged from most to least accurate. Any two treatment means not underscored by the same line are different with the probability of a Type I error equal to .05 or less.

Recalling that the smaller the W coefficient the more accurate the projection, it can be seen that Transition Analysis and Cohort-Survival Ratio Analysis produced significantly more accurate projections than did any of the other formulas investigated. This result suggests that births and tendencies of school children to advance from one grade to the next provide the most reliable predictors of future school enrollment.

Table 5 displays the results of a Scheffé test for main effects due to stratum classification. As in Table 4, any two treatment means not underscored by the same line are different with the probability of a Type I error equal to .05 or less.

Table 4.—Results of Scheffé *Post hoc* Analysis for Projection Formulas

	(1) Transition Analysis	(2) Cohort-Survival Ratio Analysis	(3) Pupil Pop. II	(4) Time Analysis	(5) Housing Analysis	(6) Pupil Pop. I	(7) Pupil-Home Unit
Mean W Coefficients	.0398	.0663	.0724	.0817	.0945	.1077	.1196
(1) .0398	*	.0265	.0326	.0419	.0547	.0679	.0798
(2) .0663		*	.0061	.0154	.0282	.0414	.0533
(3) .0724			*	.0093	.0221	.0353	.0472
(4) .0817				*	.0128	.0260	.0379
(5) .0945					*	.0132	.0251
(6) .1077						*	.0119
(7) .1196							*

$S_{(.01)} = .0363$

$S_{(.05)} = .0309$

It was significantly more difficult ($\alpha \leq .05$) to obtain accurate projections for those districts classified in Stratum A than for any other districts included in the sample. An interesting point is that accuracy of projection seemed to be directly related to the degree of stability that a given community had experienced in recent years. That is, in terms of stability, strata would be ordered D, E, C, B, A. That is exactly the same order as was derived from accuracy of projections.

Table 5.—Results of Scheffé *Post hoc* Analysis for Strata

Mean W Coefficients	Stratum (1) D .0607	(2) E .0691	(3) C .0828	(4) B .0839	(5) A .1190
(1) .0607	*	.0084	.0221	.0232	.0583
(2) .0691		*	.0137	.0148	.0499
(3) .0828			*	.0011	.0362
(4) .0839				*	.0241
(5) .1190					*

$S_{(.01)} = .0274$

$S_{(.05)} = .0231$

A third effect on accuracy of projections was educational level. More accurate projections of elementary enrollment were obtained than of secondary enrollment. This result could have been due in part to the fact that the two most accurate projection formulas studied used births as one of their predictors, enrollment in secondary school being further removed from birth than enrollment in elementary school.

The results of the initial analysis of variance also suggested a number of statistically significant interaction effects. The rejection of a null hypothesis for interaction suggests a relationship that is attributable to the joint effect of two or more variables. Therefore, the results of a Scheffé test for interaction effects is difficult to interpret since it is impossible to determine whether any differences which might arise are due to differences between main effects, differences in interaction effects, or some combination of those differences. Rather than become enmeshed in statistical tests of dubious interpretability, the data displayed in Table 6 were examined in depth to attempt to discover some relationships of practical significance.

Of the many relationships implied by the data displayed in Table 6, three stand out as having great practical significance. Of major significance was the accuracy of projections yielded by Time Analysis for districts classified in Stratum D, despite the relative lack of accuracy obtained by that methodology within all other strata. Time Analysis provided the most accurate projections for Stratum D Districts of any of the formulas investigated. This would seem to suggest that Time Analysis provides the most viable projection methodology for those districts with growth characteristics similar to those classified in Stratum D.

Also of major significance was the poor accuracy of the Cohort-Survival Ratio Method for those districts classified in Stratum B, despite that methodology's relative accuracy within all other strata. This would seem to suggest that Cohort-Survival Ratio Analysis not be used with

Table 6.—Cell and Marginal Mean W Coefficients for Projection Methodology by Stratum by Educational Level

Stratum	A		B		C		D		E		TOTAL		
Educational Level	1	2	1	2	1	2	1	2	1	2	1	2	
Projection Methodology													
Cohort-Survival Ratio	.0568	.0625	.0903	.1153	.0613	.0345	.0430	.0932	.0520	.0540	.0607	.0719	.0663
Transition Analysis	.0318	.0383	.0217	.0593	.0363	.0499	.0443	.0383	.0236	.0543	.0315	.0480	.0398
Pupil-Population I	.1200	.1774	.0704	.1659	.1362	.0756	.0794	.0662	.1327	.0534	.1074	.1077	.1076
Pupil-Population II	.0407	.1602	.0950	.0684	.0971	.0500	.0551	.0428	.0524	.0622	.0681	.0767	.0724
Pupil-Home Unit	.1944	.1823	.0782	.1304	.0729	.1992	.1156	.0925	.0782	.0529	.1079	.1315	.1196
Housing Analysis	.0615	.2870	.0601	.0608	.0432	.1349	.0743	.0522	.0892	.0821	.0657	.1234	.0945
Time Analysis	.1177	.1355	.1033	.0571	.0893	.0793	.0128	.0405	.0798	.1017	.0806	.0839	.0822
Mean	.0890	.0490	.0741	.0939	.0766	.0891	.0606	.0608	.0725	.0658	.0746	.0917	
Stratum Mean	.1190		.0840		.0828		.0607		.0692				
												Grand Mean	.0832

districts with past growth characteristics similar to those classified in Stratum B.

Finally, the fact that it was easier to project secondary enrollment than elementary enrollment in Stratum E should be noted. This phenomenon occurred despite the overall ease of projecting elementary relative to secondary enrollment. Since Stratum E contained those districts that were located in communities that experienced population decrease, this result suggests that families with older children, that is children of high school age, are more apt to remain in a declining community than are families with younger children. The relative accuracy of secondary enrollment projection for Stratum D districts, districts that were located in rather stable communities, would seem to support this hypothesis. These results are in keeping with the results of many studies on population mobility which have indicated that there is a high negative relationship between mobility and age. (We are assuming that older children have older parents). That is, the older a person is the less likely he is to migrate.

Discussion

The results of this study, although in keeping with the results of previous research by the author (13, 14) are still far too tentative. Of major interest is the replication of the results of this study in districts other than those included in the population from which the sample used in this study was drawn. This means replicating the study in districts other than those located in the state of Michigan.

The ratio and regression methods investigated in this study all have one thing in common: they are all relatively simple to apply. This is their major strength, but also their major weakness. It is a strength in that, given the applicability of a given approach to enrollment projection in a given situation, educators can apply that approach to their educational projection problems with relatively little difficulty. It is a weakness in that, given districts with certain kinds of growth characteristics, particularly districts similar in growth patterns to those classified in Stratum A, the approaches considered in this study appear to have oversimplified the estimation problem. Perhaps experimentation with more sophisticated equations incorporating various indicants of immigration and/or economic trends should be encouraged. It is quite possible that in districts which have unusual growth characteristics, a local survey that considers a larger number of demographic variables is needed in order to obtain consistently adequate projections of future school enrollment.

REFERENCES

1. Bangnee, Alfred Liu, *Estimating Future School Enrollment in Developing Countries—A Manual of Methodology*, UNESCO/United Nations Statistical Reports and Studies, Population Studies No. 40 ST/SOA/Series A, 1966.
2. Brown, Byron W., Jr.; Savage, Richard I., *Methodological Studies In Educational Attendance Prediction*, Department of Statistics, University of Minnesota, Minneapolis, September 1960.
3. Cochrane, D.; Orcutt, G. H., "Application of Least Squares Regression to Relationships Containing Auto-correlated Error Terms," *Journal of the American Statistical Association*, 44:32–61, 1949.
4. Draper, N. R.; Smith, H., *Applied Regression Analysis*, John Wiley and Sons, Inc., New York, 1966.
5. Hauser, Phillip M.; Kitagawa, Evelyn M., "On Estimation of School Population," *American Statistical Association Proceedings*, Washington, D. C., 1961.
6. Jaffe, A. J., *Handbook of Statistical Methods For Demographers*, United States Government Printing Office, Washington, D. C., 1951.
7. Johnston, J., *Econometric Methods*, McGraw-Hill Book Company, Inc., New York, 1960.
8. Kuang, Huan D., "Forecasting Future Enrollments by the Curve Fitting Techniques," *The Journal of Experimental Education*, 23:81–85, 1955.
9. MacConnell, James D., *Planning for School Buildings*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1957.
10. Scheffé, H., "A Method For Judging All Contrasts in the Analysis Of Variance," *Biometrika*, 40:87-104, 1953.
11. United States Bureau of the Census, *Current Population Reports*, Series P-25, No. 339, "Methods of Population Estimation: Part I, Illustrative Procedure of The Census Bureau's Component Method II," June 6, 1968.
12. United States National Center for Educational Statistics, *Projections of Educational Statistics To 1977–78*, U. S. Government Printing Office, Washington, D. C., 1969.
13. Webster, William J., "Comparison of the Cohort-Survival Ratio and Regression Methods for Projection of School Enrollments," *The Journal of Experimental Education*, 39:89–96, 1970.
14. Webster, William J., *The Applicability of Selected Ratio and Least-Squares Regression Analysis Techniques to the Prediction of Future Educational Attendance Patterns*, PhD dissertation, Michigan State University, East Lansing, 1969.

Forthcoming Articles

Ellis, Joseph R. and Joan L. Peterson; Northern Illinois University, DeKalb
"THE EFFECTS OF SAME SEX CLASS ORGANIZATION ON JUNIOR HIGH SCHOOL STUDENTS' ACADEMIC ACHIEVEMENT, SELF-DISCIPLINE, SELF-CONCEPT, SEX ROLE IDENTIFICATION, AND ATTITUDES TOWARD SCHOOL"

Hall, Lincoln H.; College of the Sequoias, Visalia, California
"ACADEMIC PERFORMANCES OF MATCHED GROUPS OF COMMUNITY COLLEGE AND 4-YEAR COLLEGE STUDENTS"

Among Those Present

MERL E. BONNEY is Professor of Psychology at North Texas State University, Denton. He received his BA from Willamette University, his MA from Stanford, and his PhD from Teachers College, Columbia University. He has taught in high school and in colleges in Oregon, Colorado, and California, but most of his academic career has been in Texas. He has conducted research in schools on the social development of children and youth since 1940. These studies have been reported in numerous journal articles and books.

R. DONALD BRIGGS is currently a full time counselor at Bowie State College, Maryland, where he is also teaching in the Psychology department. He completed a master's in counseling at Western Michigan and is presently working on his PhD at the University of Maryland, College Park. **DONALD TOSI** is Assistant Professor of Counseling and Guidance at Ohio State University and is involved in numerous research activities. He is widely published in educational journals and is finishing his first book in the area of counselor education. **ROSEMARY MORLEY** is a full time counselor at Kellogg Community College in Battle Creek, Michigan. She has finished work on her specialist degree at Western Michigan University and is presently involved in research and running therapy groups in the college counseling center.

THOMAS G. KENNEDY is an Assistant Professor of Education at Eastern Montana College. He received his BS and EdD from Brigham Young University. His main interest is in systems approach applied to individualized instruction. **ROBERT HUMPHREY** received his BA from Utah State University and his MA from Chapman College. He is currently enrolled in the PhD program in communications at the International University in San Diego, California, as well as being project director of COMPAS at Santa Anna College, California.

DENNIS E. NELSON is Associate Research Scientist in the Guidance Research Program of the American Institutes for Research, Palo Alto, California. He received a BS from Portland State University, an MA from Brigham Young University, and was an NDEA fellow at Stanford University where he obtained a PhD in Psychological Studies in Education. He has been with AIR since 1967 and his primary activities have been related to the development and evaluation of a comprehensive guidance system for systems of individualized education, and the development of a training program for personal problem-solving skills.

ELDON J. NULL has served as a secondary school teacher, secondary school principal, and public school superintendent in the states of South Dakota and Minnesota. He was a graduate assistant and instructor at the University of Minnesota, and received his PhD in educational administration there. Presently, he is an Assistant Professor of Educational Administration at Purdue University. He is interested in organizational theory and has carried out research designed to isolate relationships between personal variables of individuals and their perception of other persons in the school organization.

REX S. ROGERS worked on this article while Research Assistant with the Research Department of the Board of Education for the City of Toronto. He holds a BSc from the University of London, England, and is currently working as a Research Officer in the Attitude Change and Communications Research Unit of the London School of Economics and doing research toward a higher degree. **EDGAR N. WRIGHT** is Director of Research with the Board of Education for the City of Toronto. He received a BA from the University of Toronto and an MA and EdD from Teachers College, Columbia University. He taught within the Toronto school system and was an English instructor with the Toronto Teachers' College. He is also Canadian Editor of *The Canadian Lorge–Thorndike Intelligence Test*, as well as *The Canadian Cognitive Abilities Tests* and is the Research Consultant for *The Saffran Interest Inventory.*

THELMA M. URBICK received the BA and MA from Western Michigan University and the PhD from the University of Michigan. She is Assistant Professor of Counseling and Personnel at Western Michigan University, Kalamazoo, and is particularly interested in the practicum aspects of counselor education. **DOUGLAS R. GROSS** is an Assistant Professor in the Department of Counseling and Educational Psychology at Arizona State University, Tempe. He received his BS and MA degrees from Western Michigan University, Kalamazoo, and his PhD from The University of Wisconsin, Madison. He has taught at both the elementary and secondary levels and worked as both a counselor and director of a counseling program. He is currently Director of the Counseling Training Laboratory at Arizona State University and serves as Chairman of the National Membership Committee of the American Personnel and Guidance Association.

WILLIAM J. WEBSTER has his PhD from Michigan State University with emphasis in the areas of measurement, research design, and evaluation. He is currently Director of System–Wide Evaluation in the Dallas Independent School District where his responsibilities include the design and implementation of all major research and evaluation projects in the district. He was formerly a Research Scientist at the American Institute of Research in the Behavioral Sciences in Palo Alto, California.

PEARLINE P. YEATTS is presently associated with the University of Georgia, Athens, **GERALD A. STRAG** has a PhD from the University of Georgia and is an Assistant Professor at the University of Northern Iowa. He is developing methods and techniques to be used in compensatory programs with perceptually disadvantaged children.

cutive Editors

SON B. THIEDE, Professor of Curriculum and Instruction, University of Wisconsin, Madison, Chairman.

NE OTTO, Professor of Curriculum and Instruction, The ersity of Wisconsin, Madison.

JOSEPH LINS, Director of Research, Wisconsin Coordinat- Committee for Higher Education, Madison, Wisconsin.

nsulting Editors

Terms Expire December 31, 1971

RENCE P. BLUM, Professor of Educational Psychology, ersity of Wisconsin—Milwaukee.

GLAS R. GROSS, Arizona State University, Tempe.

LLIAM B. McCOLLY, Professor of English, University of th Carolina, Columbia.

LLARD E. NORTH, Director of Research, Central Missouri te College, Warrensburg. Field News Editor.

N R. PALMER, Professor of Curriculum and Instruction History, The University of Wisconsin, Madison.

N E. STECKLEIN, Bureau of Institutional Research, Uni- ity of Minnesota, Minneapolis.

LIA R. VANE, Professor of Psychology, Hofstra University, npstead, Long Island, New York.

Terms Expire December 31, 1972

RMAN J. BOYAN, Dean, Graduate School of Education, ersity of California, Santa Barbara.

BERT D. BOYD, Associate Professor of Curriculum and truction, University of Wisconsin, Madison.

BERT CLASEN, Instructional Research Lab., University of consin, Madison. Book Review Editor.

RROLL L. GONZO, Assistant Professor of Music Education, ersity of Illinois, Urbana.

B W. JERROLDS, Assistant Professor of Education, Univer- y of Georgia, Athens.

RL KOENKE, Assistant Professor, College of Education, ersity of Illinois, Urbana.

ROLD E. MITZEL, Assistant Dean for Research, College of cation, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

GER W. BARDWELL, Superintendent of Schools, Ithaca y School District, New York.

LORES DURKIN, Professor of Elementary Education, Uni- ity of Illinois, Urbana.

VER L. HURLEY, Associate Professor of Special Education, hiva University, New York.

ORGE S. JURATA, Coordinator of Reading, Jefferson County ools, Colorado.

MES D. RATHS, Professor of Education and Director, reau of Educational Research and Field Services, University Maryland, College Park.

JAY SAMUELS, Associate Professor of Educational Psy- ology, University of Minnesota, Minneapolis.

RMAN R. STEWART, Professor, Department of Counseling, sonnel Services and Educational Psychology, Michigan State versity, East Lansing.

itors Emeriti

RRY J. BAKER, 2241 Q Via Puerta, Laguna Hills, Cali- nia 92653.

WLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 757.

RL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

ARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, lifornia 92802.

M. JORDAN, Professor Emeritus, School of Education, Uni- sity of North Carolina, Chapel Hill, North Carolina 27514.

LFRED S. LEWERENZ, Director Emeritus of Evaluation and search, 4107 Holly Knoll Drive, Los Angeles, California 027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, lorida 33146.

ESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

368
13 · 9 · 71

Bureau Ed. P Y

51st Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 **MAY-JUNE 1971** **NUMBER 9**

ARTICLES

State of the Art of Teacher Education—*Lindley J. Stiles* --- 387

Testing for Number Readiness: Application of the Piagetian Theory of the Child's Development of the Concept of Number—*Robert Williams* --- 394

Field Dependence—Field Independence in 5- to 10-Year-Old Children—*Gordon S. Bigelow* --- 397

Characteristics of Students Who Make Accurate and Inaccurate Self-Predictions of College Achievement—*Karl E. Keefer* --- 401

Faculty Allocation of Effort and Student Course Evaluations—*Claude W. Grant* --- 405

The Effect of Selected Demographic Factors on Prediction of GPA by Male Freshman Students—*Richard F. Sebok* --- 411

Sex Differences in the School Gains of Migrant Children—*Sushila Singhal; Priscilla H. Crago* --- 417

The Effect of Grade Contracts on Student Performance—*William A. Poppen; Charles L. Thompson* --- 420

Sociometric Predictors of the Self Concept Among Fifth and Sixth Grade Children—*Bert O. Richmond; William F. White* --- 425

DEPARTMENTS

Among Those Present --- 432

Book Reviews—*Robert E. Clasen* --- 424

Field News—*Willard E. North* --- 430

Forthcoming Articles --- 419

It Says in The JER --- 386

New Books from the Publishers --- 416

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, Supervisor of Editorial Services.

It Says in The JER...

"TEACHER EDUCATION TODAY, says Stiles, must be described as confused and insecure." Teacher education must learn to shun all orthodoxies and make revolution a continuing tradition to develop teachers for the times—past, present, and future (p. 387).

KINDERGARTEN MAY BE too early for addition and subtraction, discovered Williams. He tested ninety-six kindergarten children with Piaget's 3-stage theory of the child's attainment of the concept of number with a number readiness criterion and found only eight at the stage of attainment which is inclusive of operational ability (p. 394).

SOCIAL CLASS may affect criterion performance, but age is the best predictor of criterion performance as significant differences proceeding from relative field-dependence to greater field independence with increasing age were found by Bigelow. He found no significant similarities between field-dependence-independence and verbal intelligence within three of four specific ages he investigated (p. 397).

CAN COLLEGE STUDENTS in the top quartile predict their own grades better than students in the bottom quartile? Keefer found that those who are accurate predictors of their own achievement tend to achieve at a significantly higher level than those who are inaccurate. He found also that upperclassmen are better predictors than freshmen (p. 401).

FACULTY RANK was not related to student evaluations of courses on any of five course evaluation scales and in general there was little relationship between faculty allocations of time and student evaluation of courses. Grant noted, however, that students in large enrollment courses considered faculty more "prepared" than did students in small enrollment courses (p. 405).

ECONOMIC FACTORS were significant in student prediction of college academic success in Sebok's study. Students from low economic backgrounds predicted lower grades. He says it may be hypothesized that prediction would be affected by freedom from economic worry and financial aid could enhance their opportunities for academic achievement in college (p. 411).

ARE MIGRANT children's school gains affected adversely either by teachers' discriminatory behavior in favor of girls or by family structure? Singhal and Crago found no indication of this in their study of school gains of migrant children. There were achievement differences at various elementary levels between sexes (p. 417).

GRADE CONTRACTS were studied by Poppen and Thompson. They found there were no significant differences between the grade contract approach and the traditional approach to instruction at the college level (p. 420).

THE COOPERSMITH Self-Esteem Inventory and a semantic differential were used by Richmond and White to assess children's attitudes toward themselves and to rate peer concepts in the classroom. The five factors they discuss help to identify characteristics of the children and peer concepts of each individual type of child (p. 425).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Editorial/Abstract

The State of the Art: Teacher Education

◆ In the past four decades, teacher education has responded to the times: the search for self fulfillment of the 1930's Progressive Education reflected an answer to the problems of social turmoil during an economic depression; post World War II economic and technological advances prompted an emphasis on academic knowledge with little stress on skills promoting learning; by the 1960's the trend was to increase clinical experiences and to de-emphasize intellectual study of education and its process.

With the sixties came a counterrevolution, stemming from social, economic, and political conditions. Minority groups questioned the ethnic and racial compatibility of teachers; unemployment and inflation deflated the status of teaching; students demanded an active place in their own learning and relevancy in teaching. In the future, teacher education can be expected to change with the forces that play upon it, if traditional patterns of response continue.

Because teacher education is more art than science, more political than professional, and more relevant socially than disciplines which are purely academic, conflict is characteristic. It seems unlikely that educating educators will ever be free of controversy, as social and political forces will continue to challenge any models of teacher education in existence.

The challenge for teacher education is to live with the controversy rather than try to counteract it. The quest is to make teacher education relevant, and the state of the art is in its infancy.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

State of the Art of Teacher Education

LINDLEY J. STILES
Northwestern University

NO PROFESSIONAL field is as consistently controversial as teacher education. The most predictable characteristic about the art of teacher education is that whatever exists today will be challenged tomorrow. The pattern is cyclical, moving through successive stages of indictment, innovation, dissemination of new orthodoxies, defense of the status quo—and then revolution all over again. In the continuing debates, everyone gets into the act, students, parents, politicians, academic scholars, as well as teachers and administrators themselves, and of course, specialists in teacher education. It is a carrousel of conflict that returns again and again to the same issues and the same arguments as well as to many of the same proposed solutions.

Conflict is characteristic of teacher education expressly because it is more art than science, more political than professional, and more socially relevant than purely academic. The scientific movement in education is still in an embryonic state. It has yet to unravel the mysteries of human learning and its modification to the point that reliable directives can be given to teaching. In the absence of definitive research results, passion and politics play disproportionate roles in establishing directions, priorities, and practices in teacher education. Fads and fashions, the theories of popular leaders, proposals sponsored by professional organizations, programs of prestigious institutions, formulas implanted by philanthropic foundations, and more recently, the federal government, as well as public and professional acceptance, comprise the constant cross-currents that keep teacher education going in circles.

The dream that ultimately teacher education will emerge from its wilderness of controversy—led either by research or the wisdom of a renowned leader, as Flexner did for medical education and Conant (3) recently tried for teacher education—seems unlikely ever to be fulfilled. Even when research evidence becomes available to help define what teacher education should be—as instructive as it will be—social and political forces will continue to challenge whatever models of teacher education exist at a given time. Nor can the wisdom of great leaders, or of committees of educators, much less the proposals of popular critics, no matter how much they are respected, lay down a prescription for teacher education that will stand the test of time. Science cannot set the goals for the society, its schools, and hence, for teacher education. Proposed and accepted philosophies and programs are fragile in the face of modifying social, political, and economic conditions. The persistently changing needs and aspirations of people and the negotiated compromises that are evolved among competing demands prescribe what teacher education will be. To understand teacher education, then, at a given point in time, one must be sensitive to the social milieu in which it exists and appreciate the dominant forces that play upon it. It helps, too, to perceive the stage in the cycle of response that teacher education has reached.

Teacher Education in Revolution

Teacher education in the United States has been moving into a new revolution over the past decade. The current renaissance follows a preceding period of reform, during the 1930's, when teacher education faced the problem of adjusting to the social turmoil of a major economic depression that made schooling more of an escape from reality than a preparation for life. Progressive education was the schools' answer to a disillusioned age. Children and youth were taught to look within themselves for fulfillment. Human relations were stressed as ends in themselves. Creativity was encouraged. School was expected to be fun and satisfying as a way of living rather than as rigorous preparation for something to come. Institutions preparing teachers were challenged then, as now, to prepare teachers for the times.[1] During the 1940's and 1950's the new orthodoxy for teacher education, in philosophy if not always in practice, became widely disseminated. By mid-century, leaders in teacher education felt so confident that they had, at long last, found the way that they set about building defenses for the status quo to assure that the pattern would be perpetuated.

As teacher education was settling in for its anticipated long mid-century sleep, a new revolution was brewing. A period of rapid economic and technological advancement following World War II abruptly changed the signals for education. Schooling was no longer viewed, as had been the case during the depression years, as wholesome experience for the young in a society that offered little incentive for academic achievement; it was now the instrument to prepare skilled workers, and particularly high level scientific personnel, for an expanding economy that had an insatiable appetite for people with marketable skills and highly organized brain power. Academic knowledge, particularly in fields such as science and mathematics, was viewed as a national necessity. Such political and economic changes altered the demands on teachers. They were challenged to quit "playing around" with students and to get busy teaching them the academic knowledge and intellectual skills that the market demanded. Teacher education was indicted for failure to provide the kind of rigorous academic preparation that such teaching required. The demand was for more emphasis on subject matter in the teacher's preparation. To make room for such additional content, it was argued that the time given to professional preparation be reduced, particularly in teaching methods. An underlying assumption was that teachers did not need much knowledge about or skill in promoting learning since the rewards offered by an affluent society were sufficient to motivate students. Poor quality in some education courses—due partly to lack of reliable research evidence—added impetus to the argument. The criticisms and proposed solutions pointed to an earlier tradition when academic preparation alone and on-the-job apprenticeship training were considered sufficient preparation for teaching.

The mid-century orthodoxy did not give way easily to the demands for reform. Throughout the 1950's, it withstood extensive public criticisms and resisted new innovations, inaugurated principally by the Ford Foundation (14). Ironically, although the pattern of teacher education had grown out of the era of progressive education, in which creativity, individuality, and reflective thinking were prized, and despite the fact that it gave lip-service to progressive ideals, in practice it tended to evolve into a rigid, monolithic, sterile, and unchallenging process. The plan was devised and controlled by professional educationists, with little involvement of academic scholars or teachers and administrators in elementary and secondary schools. Political manipulations by the education establishment had given the system the protection of state certification laws and a newly created National Council for Accreditation for Teacher Education (NCATE). Teachers and administrators tolerated and supported the system because they had come to identify professional preparation as a key element in their drive to achieve a higher level of status for themselves. That the program had many weaknesses, as critics generously pointed out, was recognized by the educational establishment. The prevailing spirit, however, was that a "united profession" was worth the costs of uniformity and other deficiencies that existed. Thus, loyalty to the established orthodoxy was demanded from all educational personnel and criticisms, either from within or without, were interpreted as subversive efforts to undermine the profession itself.

Despite the determined defense of the status quo, public pressures increased to the point that by the early 1960's it was clear that a new revolution was brewing. Teacher educators began to talk about reforms and to advise each other about the changes that would be desirable (2). Teacher preparing institutions began to adopt modifications in their programs, many of which had been stimulated by the new models they had been resisting. Cracks in the walls of the orthodoxy began appearing as different states—California, Florida, Hawaii, New York, Washington, and Wisconsin, to illustrate—began to make changes in certification prescriptions. The National Commission for Teacher Education and Professional Standards, NEA, began advocating a program-approval rather than a credit-counting approach to teacher certification.

A dramatic break in the defenses of the orthodoxy came with the now historic confrontation between The University of Wisconsin and the NCATE. The University refused, as ordered by the Council, to make policy for teacher education the exclusive prerogative of educationists, to reduce its alternate programs of preparation to one pattern, or to require uniform points of entry and patterns of progression through its various programs of teacher education. The NCATE, committed as it was to defending the status quo of a monolithic model and to preserving the control of teacher education "in our hands" as one dean of education candidly put it—threatened to withdraw the institution's accreditation. In the ensuing conflict, in which academic and public opinion supported The University of Wisconsin and the educational establishment lined up behind the NCATE, it was clear that strongly defended traditions in teacher education were at stake. As history records, Wisconsin was the victor in the controversy. The NCATE was forced to reorganize itself and to modify its standards to allow for greater flexibility, more institutional autonomy, and interdisciplinary control of teacher education.

Challenges of the Sixties

In the face of criticisms of education courses, the response was to increase the amount of clinical experiences and to de-emphasize the intellectual study of education and its processes. A key force in the shift of emphasis from an academic to an apprenticeship approach was support of the Ford Foundation for a new type of master of arts degree in teaching (MAT), first on a statewide basis in Arkansas and then in selected prestigious institutions. The MAT idea, which had been pioneered originally by Harvard University, aimed to involve liberal arts and education professors as well as colleges and elementary and secondary schools in common partnership in the preparation of teachers. By postponing professional preparation until the fifth year, these programs achieved several objectives: they virtually eliminated the competition for the student's undergraduate time that pitted academic and educationist professors against each other; they opened the door for graduates of liberal arts colleges to prepare to teach in elementary and secondary schools; and they provided a full-time internship for a semester or year. A central thrust was to relate the study of theory and knowledge about education to the internship which made the process of learning to teach more inductive.

In practice, MAT programs vary from carefully planned innovations to nothing more than traditional courses in education and student teaching. A modification of the approach is the intern-in-team plan in Wisconsin, also supported initially by the Ford Foundation, which shifts responsibility for instruction in methods of teaching to clinical supervisors in elementary and secondary schools and provides for the intern to be a member of an instructional team while in training.[2]

Another version of a MAT program which challenged teacher education during the 1960's was the Cardozo Plan (9) in Washington, D.C., a forerunner of the National Teacher Corps Program. It aimed specifically to prepare young people with experience in the Peace Corps and other volunteer organizations to teach inner-city (black) children. The process was primarily tutorial and clinical with interns working in small teams under the supervision of expert teachers. The Teacher Corps Program, established by the Federal Congress in 1965, followed the basic pattern of the Cardozo Plan with teams of volunteers working under the leadership of master teachers to improve the learning of minority group students. Preparation for teaching, which is accomplished by tutorial and seminar techniques, is a by-product.

A contrasting approach to teacher education, which also moved with the impetus toward more emphasis on practical experience, is the tutorial-clinical approach pioneered at Northwestern University (6). Supported during its pilot stage by the Carnegie Foundation, this program rejects the idea that professional preparation should be postponed to the fifth year of the student's collegiate study. Instead, it involves the prospective teacher in firsthand experiences in schools and the intellectual analysis of professional problems throughout the 4 years of undergraduate study. A feature of the tutorial-clinical approach is the extent to which preparation for teaching is individualized with each student designing his or her own professional preparation to meet individual needs and preference with the help of faculty tutors and clinical professors in the schools.

An innovation in recent years that has brought new emphasis on teaching strategies is the idea of micro-teaching, pioneered at Stanford University (1). This procedure provides for prospective teachers to refine skills of instruction, such as the techniques of asking questions or posing problems, in small group situations and in short practice exercises of a few minutes duration each. Micro-teaching reemphasizes the fact that instructional techniques make a difference in learning and documents that intellectual analysis as well as experience is essential to professional development.

Throughout the past decade little attention was given to improving the liberal or general education of teachers. Action has been taken, however, to reorganizing the content of the teaching fields, a movement led by academic scholars. Called the "new curriculum," the objective has been to identify the theory and structure of a discipline. Thinking or inductive approaches to study have been favored.

Of the various influences on teacher education that are developing, the growing emphasis on basic research may have the most lasting effect. A landmark study was one by Ryans which identified teacher characteristics as a base for refined investigations of each (7). The publication of an extensive analysis of research on teacher education with proposals for new designs added to the foundation for the scientific investigation of the field (5). Significant emphases in basic research have centered on the study of the classroom behaviors of teachers, the effectiveness of specific instructional tactics, and various other factors in teaching that make a difference in student learning (4). Research on programmed instruction, related to the new uses of technology in education, has added insights into how learning most efficiently is promoted (13). Slowly the stockpiles of knowledge about teaching and learning are growing, giving reassurance of the basic truth that the professional practice of teaching needs a sound base of professional theory and knowledge as well as accumulated practice—a point that is obscured by the current focus on firsthand experience.

Counterrevolutions

Before the revolution of the 1960's had become widespread, a number of counterrevolutions were forming. Changes in social, economic, and political conditions were forcing new and contradictory demands on schools and teachers. Whereas, at mid-century the challenge was for greater academic proficiency, to produce high level brain power and skilled technicians for a scientific age, it soon became apparent that for many students—principally the non-white minorities—such a goal was not realistic. Consequently, a new type of criticism was raised against teacher education; namely, that it was developing teachers to serve only the white, middle class, suburban-type students. A related charge was that the teachers being produced were using instructional procedures and managing the educational resources to favor the fortunate rather than to help the disadvantaged. Teachers for the "real world" of the ghetto were demanded—a mandate that runs counter to the emphasis on developing the intellectually talented that the reformers of the 1950's had promoted. Thus, teacher education is faced with conflicting demands—a revolution and a counterrevolution, each posing its own vested goals and together generating confusion as to purposes and procedures.

Recent changes in social and economic factors have further confused the mandate for teacher education in the immediate years ahead. Minority groups have raised the question of ethnic and racial compatibility of teachers. In the extreme, the suggestion is that only the black can teach the black, or the brown the brown. More moderate interpretations call for teachers who understand and are empathetic to particular ethnic or racial groups and who are able to win their confidence. Interpreting the meaning of such ideas for teacher education is further complicated by school integration actions that assign teachers to schools on quota bases. Confusing as it is, the concern for producing teachers who can make learning relevant to particular kinds of students is challenging the long-held assumption that teachers, as are building bricks, are interchangeable from school to school, from grade to grade, and from one cultural group to another.

Another type of counterrevolution that may change teacher education in the years ahead is the economic trend toward high unemployment and continuing inflation that brings depression conditions to so many people. When professionals with doctors degrees are walking the streets looking for work, as was the case during the depression of the 1930's, and is becoming true again today, the mandate that schools develop high level intellectual skills—the central impetus for the revolutionary pressures of the past 20 years—loses its potency as motivation for learning. Already, the suggestion has been made that schools return to the philosophy and practices of the progressive education era (8). The argument advanced is familiar: that schooling should be fun and rewarding for its own sake rather than as preparation for the future.

Teacher education is affected, of course, by the student revolt movement, another type of counterrevolution. Obscured by the militant actions of extreme radicals who seek to overthrow the society, and apart from the protests against the war, racial injustices, and environmental pollution, are growing demands of students that teaching be improved. Young people believe that they should be actively involved in their own learning; they want to share with teachers the policy-making and planning—to become full partners in the teaching-learning enterprise. A second type of demand that is familiar to all is that teaching be relevant. Here the concern is for pertinency of instructional strategies as well as in terms of content selection. Activism is a goal that students endorse; they believe that teaching should make a difference, that learning by firsthand experiences is vital to changed behavior; that teachers and students should translate theories and principles into practice. Some of the student protests against teaching are directed against the poor rapport that often exists between teachers and students. They want to know teachers as human beings, to trade ideas and feelings with them in intimate situations.

The mushrooming of current counterrevolutionary movements with each posing for teacher education its own set of special mandates illustrates why no model or pattern of teacher education is likely to be more than a temporary orthodoxy. The quest will always be to prepare "teachers for the times"; the problem will be to understand the times and to change with them. But the times are not monolithic; they vary with different cultural, economic, and political groups.

Accent on Innovation—Open for Research

Teacher education has moved into the 1970's generally committed to innovation and with an openness for research that is perhaps unprecedented.[3] Out of the controversies and pioneer programs of the decade past, and with the restrictive prescriptions of certification and accreditation reduced, there has evolved a general receptiveness to change.

Supporting the new commitments to innovation and research are certain policy and philosophical benchmarks of agreement—which, of course, may be only temporary. Foremost among these is a general acceptance of the idea that professional skills and knowledges essential for successful teaching are learned rather than inherited. Academic scholars, many of whom have long championed the latter view in their criticisms of teacher education have been forced by recent attacks on their own pedagogy into partnerships with educationists to learn how to improve instruction.[4]

Agreements now seem to prevail that teacher education should be interdisciplinary in character with policies as well as programs being developed and carried out by both academic and professional personnel (10, 12). The "all-university" approach, however, is as yet more of an ideal than a reality in many institutions. In practice, educationists still exercise controlling power to determine how teachers will be prepared.

A commitment exists to involve elementary and secondary school systems in full partnership with institutions of higher learning for teacher education purposes. The belief that school systems, and the practicing professionals in them, should take more or perhaps full responsibility for providing the clinical aspects of the teacher's preparation is gaining momentum. Not only are teachers becoming interested in sharing responsibility for preparing members of the profession, many are beginning to insist that they have a major role in judging qualifications for practice (11).

The emphasis on clinical experiences of the past decade is continuing. New innovations focus on action programs that build professional knowledge around clinical training. A new dimension is the recognition that the kind of clinical situation makes a difference in the kind of teacher produced.[5] Thus, it is no longer assumed that a teacher for the inner-city schools, for example, can be prepared in a campus laboratory school or in a suburban situation. Particular attention is being given to clinical experiences in ghetto schools where improvement of the learning of culturally different students and teacher development are intertwined in common cause.

Emerging Priorities

To predict the direction of teacher education is a precarious venture in view of the multiple revolutionary forces that now prevail. It is safe to assume that no group of critics will propose long-lasting prescriptions, nor are any of the innovative models likely to become definitive. In view of the traditional cycles of response to the social milieu of given times, teacher education can be expected to change with the forces that play upon it. Thus, the current revolutions can be predicted to produce their own orthodoxies which in turn will be challenged by new demands for reform.

The emerging priorities in teacher education in the immediate years ahead appear to include the following:

1. *There will be an increased emphasis on research, financed most likely by the federal government, but hopefully also by the philanthropic foundations, and state governments, as well as by schools and colleges. The priority will be for developmental research that aims immediately to establish new models for teacher education that will produce teachers with particular talents and commitments.*

2. *While research knowledge is accumulating, it must be assumed that teaching will continue to be a personal endeavor, with each practitioner giving expression to his or her unique perceptions and talents. In this connection, a trend can be anticipated toward more student involvement in the process of instruction to the point that traditional procedures for preplanned lessons will give way to more spontaneous and creative teacher-student cooperation.*

3. *Teaching will continue to move toward being viewed as a specialized function with individuals being prepared to instruct particular kinds of students, such as representatives of given ethnic or racial cultures, as well as given skill or subject fields. Performance of specialized functions in instructional teams will be a related emphasis.*

4. *The process of learning how to prepare teachers for particular kinds of students will be intertwined into efforts to discover how to promote learning for such students. Thus, teacher education will be more school than campus centered. As knowledge is refined in operational situations it will become imbedded in formal teacher education programs.*

5. *There will be a renewal of concern for the affective influence of the teacher that will bear on how teachers are prepared.*

6. *Current efforts growing out of the use of new instructional technology to make learning "teacher proof" will give way to a broader view that human beings have unique contributions to make to instruction that are beyond the scope of automated technology. Research will aim to differentiate the types of higher emotional and intellectual contributions that teachers can make as contrasted to the more routine procedures that programmed instruction can monitor. Eventually a workable compromise between human and technological resources will be established.*

7. *More attention will be given to the human relations skills and sociological understandings of the teacher.*

8. *As awareness evolves that clinical experiences alone are insufficient to produce the professional teacher, we can expect a renewal of attention to the kinds of theory and knowledge that are pertinent to effective teaching and professional responsibility.*

9. *Interdisciplinary and school and college partnerships will be expanded with responsibility being divided in accordance with resources and capabilities. Academic professors in colleges and universities will take more*

interest in teacher education as a means of improving their own instruction as well as helping to develop teachers for elementary and secondary schools.

10. Members of the teaching professions will take greater interest in teacher education, seeking to influence policies as well as to control the licensing of personnel, in addition to supervising the clinical practice of prospective teachers.

11. If the oversupply of teachers continues, pressures can be anticipated to extend the period of preparation for initial teaching to a minimum of 5 years. A related influence will be to curtail the use of instructional aids and Teacher Corps trainees.

12. Differentiated staffing will help to make teaching a career profession, with commensurate rewards and prestige, which will increase the demand for in-service programs that contribute to the development of high level instructional competence.

The Futile Search for Orthodoxy

Teacher education today must be described as confused and insecure, beset by multiple revolutionary forces, and madly searching for a new orthodoxy. With confidence in professional theory and knowledge shaken, it has returned to the apprenticeship in search of relevance and to escape from criticisms. In desperation, innovative models are being conceived and put into operation, trying to find a new pattern to replace the discredited old. The commitment is to find a formula, or formulas, that will mass produce teachers into predetermined professional products. Innovation is the tactic employed which follows an easily recognized political formula: when criticisms generate, try something different. The results are piecemeal; the process is futile.

Progress in teacher education will come with research that refines and tests theory, studies and solves specific problems, and projects deeper insights about learning and human development and their modification—rather than new orthodoxies. A need is to examine instruction itself—in terms of its artistic and scientific elements, its human and technological components, its varied impact on different kinds of students—in contrast to the concern for form (models) of the program of teacher education overall.

To develop teachers for the times, present and future in all their pluralities, teacher education must learn to shun all orthodoxies and to make revolution a continuing tradition. The challenge is to live with controversy rather than try to counteract it. The persistent quest is to make teacher education relevant, to social conditions as well as to refined knowledge, and especially to the individuals who are preparing to teach. The state of the art is in its infancy.

FOOTNOTES

1. Then the challenge was voiced by the Commission on Teacher Education of the American Council on Education in a report called *Teachers of Our Times*, American Council on Education, Washington, D. C., 1944. Today a similar appeal comes from the American Association of Colleges for Teacher Education under the title, *Teachers for the Real World*, AACTE, Washington, D. C., 1969.
2. Many of the elements of this plan are endorsed by John Goodlad in a proposal he recently made for "The Reconstruction of Teacher Education," *Teachers College Record*, pp. 61–72, September 1970.
3. In 1971, several hundred institutions were developing new models for teacher education with support from the Education Professions Development Act of the federal government. Others are engaged in research and experimentation of various kinds.
4. Indicative of this trend was the establishment, in 1969, of a National Center for the Teaching Professions at Northwestern University, the primary aim of which is to help all professional and academic fields to improve their instruction. Basic support comes from the Kellogg Foundation with the Center being administered by the School of Education.
5. As illustrated in the models for training teachers in the Education Professions Development Act programs. See: Education Professions Development Act *Handbook*, U. S. Government Printing Office, Washington, D. C., 1969, p. vii.

REFERENCES

1. Allen, Dwight W., *Micro-Teaching: A Description*, Stanford University, Palo Alto, California, 1967.
2. American Association of Colleges for Teacher Education, *A Decade of Thought on Teacher Education*, The Association, Washington, D. C., 1969.
3. Conant, James B., *The Education of American Teachers*, McGraw Hill, New York, 1963.
4. Flanders, Ned A., "Teacher Effectiveness," *Encyclopedia of Educational Research*, The Macmillan Company, New York, 1969, pp. 1423–1434.
5. Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand McNally, Chicago, Illinois, 1963.
6. Hazard, William R., *The Tutorial and Clinical Program of Teacher Education*, Northwestern University Press, Evanston, Illinois, 1967.
7. Ryans, David G., *Characteristics of Teachers*, American Council on Education, Washington, D. C., 1960.
8. Silberman, Charles E., *Crisis in the Classroom*, Random House, New York, 1970.
9. Stiles, Lindley J., *Cardozo Project in Urban Teaching: Evaluation and Recommendations*, report by the project, February 1967.
10. Stiles, Lindley J., "Interdisciplinary Accountability for Teacher Education," *Journal of Higher Education*, pp. 23–31, January 1968.
11. Stiles, Lindley J., *Teacher Certification and Preparation in Massachusetts*, Massachusetts Advisory Council on Education, Boston, 1968.
12. Stiles, Lindley J.; Carver, Fred D., "Who Makes Policy for Teacher Education?" *Teachers College Record*, pp. 203–211, December 1967.
13. Stolurow, Lawrence M., "Programmed Instruction and Teaching Machines," in Rossi, Peter H.; Biddle, Bruce J. (eds.), *The New Media and Education*, Doubleday and Company, Inc., Garden City, New York, 1967, pp. 136–192.
14. Woodring, Paul, *Investment in Innovation: An Historical Appraisal of the Fund for the Advancement of Education*, Little, Brown, and Company, Boston, Massachusetts, 1970.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Testing for Number Readiness: Application of the Piagetian Theory of the Child's Development of the Concept of Number

ROBERT WILLIAMS, University of Dayton

ABSTRACT

Piaget's 3-stage theory of the child's attainment of the concept of number was applied as a number readiness criterion to ninety-six kindergarten children. Two children were found to be at the initial stage of number development, eighty-six at the intermediate stage, and eight at the stage of attainment of number concept which is inclusive of operational ability. Results were applied to mathematics curriculum goals on the kindergarten level. Curriculum goals for mathematics on this level should not include operations (addition, subtraction). Socialization to number symbols may be established as a goal at this level, but it has nothing to do with the concept of number and does not appear to enhance attainment of the concept of number.

PIAGET'S MODEL of the development of the number concept identifies three stages, through which each child must progress: (1) an initial stage, (2) an intermediate stage, and (3) a stage of attainment of number concept. According to Piagetian theory (1:260–261,2,4,5), the attainment of the concept of number (Stage III) is demonstrated when a child (6½ to 7 years) is able to equalize a collection of elements (e.g., 8 disks) 1 inch apart, in a row with another collection of elements by bi-univocal correspondence between their terms. Piaget further maintains that there must be perceived equivalence of the collections after the density of the elements composing one of the collections is modified in such a way that the elements of one row no longer appear directly opposite those of the other. Before the attainment of the concept of number, the child's response takes one of two similar but incomplete forms. In the intermediate stage (Stage II), the child (5 to 6 years) makes a one-to-one correspondence of the two collections as a basis of equivalence; but after a modification of the density of the elements, the child loses equivalence of the collections. During the initial stage (Stage I), the child (4 to 5 years) responds to the elements of the model row by constructing a row equal to the length of the model but of unequal elemental distribution and composition.

Before discussing application, this model will be identified among Piaget's sequential-developmental stages. The three sequential stages toward the attainment of the concept of number seem to be within the transitional phase at the conclusion of the preconceptual phase, the whole of the intuitive thought stage, and the initial phase of the stage of concrete operations. According to Piaget, this model of the development of the number concept cannot be effectively applied before the age of 4 or the stage of intuitive thought, at which time short experiments involving the manipulation of experimental objects and meaningful conversation about the activity may be carried on (2:162, 5:129).

Procedure

An incidental sample of ninety-six children ages 5:6:26 to 7:1:0 (years:months:days) of Weller Elementary School, Centerville, Ohio, was interviewed with the intention of: (1) establishing whether Piagetian stages actually existed within the group; (2) identifying the stage of each child at the date of testing, and (3) relating the results to mathematics goals at the kindergarten level. During the interviews each child was presented with a model row composed of eight disks of the same size and color.

○ ○ ○ ○ ○ ○ ○ ○ (Model)

The child was then asked to place down as many disks from his collection (which were the same size as those of the model but of a different color) as there were in the model. A child at the initial stage (I) might respond in the following manner:

○ ○ ○ ○ ○ ○ ○ ○ (Model)
○○○○○○○○○○○○○○○ (Response)

The child has constructed a row equal in length but of unequal elements and composition; yet he believes that he has established equivalence between the two collections. At the intermediate stage (II) a child responds by establishing a one-to-one correspondence of the elements of both collections.

○ ○ ○ ○ ○ ○ ○ ○ (Model)
○ ○ ○ ○ ○ ○ ○ ○ (Response)

The child will affirm the equivalence of the collections before a modification of the density. The test of the equivalence of this one-to-one correspondence comes when the density of the elements of one of the rows is modified:

o o o o o o o o
oooooooo

After modification of the density of the elements, a child of the intermediate stage will affirm a lack of equivalence of the collections; however, at the stage (III) of attainment of the concept of number, the child maintains equivalence of the collections.

Every child was asked to make a manual response. The initial stage was identified from this manual response. Children making a one-to-one correspondence were interrogated after a modification of the density and a destruction of perceptual correspondence. The verbal response was the determinant in identifying the intermediate and attainment stages (2:26).

Results and Discussion

All Ss could be adequately categorized by Piaget's model of number concept development. Two children (male, 5:10:14 and female, 6:3:23) were found to be at the initial stage. Neither corresponded with the hypothesized age of the initial stage (4 to 5 years). This lack of correspondence is not a critical element. Piaget cautions against an over-literal identification of stage and age. His findings give only a rough estimate at best of the mean age at which various stages are achieved (2:20). Eighty-six of ninety-six were in the intermediate stage. Nine of the eighty-six rendered a verbal count of the number of elements composing the rows before perceptual correspondence was destroyed. As Piaget has maintained, concerning children of the intermediate stage, correspondence and the ability to count the elements of the rows were of no aid to their maintenance of equivalence after perceptual correspondence was destroyed (5:144).

Neither the initial nor intermediate stages manifest the presence of operational ability or the attainment of the concept of number. The child of the initial stage does not differentiate the disks which compose the row from the row itself and his respose is regulated by what Piaget refers to as global comparison (4:86–7). The responses of children of the intermediate stage reveal that they take the immediate conditions of the model as the final and conclusive reality. The children seem to have given their attention to the end results of the transformation of the model (modification of the density of one of the collections) rather than to the transformation itself (2:157). Concerning centration on the end result of transformation Flavell states,

> *A single, isolated cognition of this sort, with little or no systematic reference to other cognitions past or potential, is the hallmark of the preoperational child* (2:167).

Eight of the children (5:11:3 to 6:8:28) were able to maintain the equivalence of the two rows after perceptual correspondence was destroyed. These children were judged to have attained the concept of number and manifested the presence of operational ability which within the Piagetian system is action or a system of bodily movements which have become internalized in the form of thought activities (1:xvi, 155–56, 167). Flavell in his explanation of Piagetian cognitive operations states,

> *A useful rule of thumb, one Piaget has used (5:32–33), is to say that all the actions implied in common mathematical symbols like +, –, ×, ÷, =, <, >, etc., belong to, but do not exhaust the domain of what he (Piaget) terms intellectual operation* (2:166).

Children of the stage of attainment do not center their attention only on the end result of transformation but also upon the transformation. The transformation that now concerns the child is merely one example of a whole system of possible transformations. Each transformation is now seen in relation to another transformation which can negate it. Here we observe a maturation of the coordination of reversible cognitive actions which distinguish the preoperational child from the child in the stage of concrete operations (2:167).

Relating the results of this study to kindergarten mathematics' goals is germane. For children identified in the first two stages we cannot have operational objectives (addition and subtraction) in mathematics. It is not realistic to have the learning of addition and subtraction as objectives. Addition and subtraction involve operations which are not congruent with the cognitive structures of the children. At this time the only realistic goal that may be considered is to give attention to the socialization of the children to the culturally accepted symbols for numbers, 1, 2, 3, 4, 5, etc. A reservation is that socialization concludes with the number 10 unless the child shows a desire to go further. It is sufficient for the purposes of socialization that children be given only the very basic numbers. What occurs in socialization has nothing to do with the concept of number. That this group of children was socialized to number symbols was verified through a study of the results of ninety-four of ninety-six of the children's Metropolitan Readiness Tests, Form A, Test 5 Numbers (MRT,FA,T5N) (3) and especially the questions from that test which

Table 1.—Results of Questions from the Metropolitan Readiness Tests, Form A, Test 5 Numbers Which Seem to Reveal the Rote Knowledge the Children had of Certain Social-Cultural Number Symbols[1]

	Correct N	Correct %	% of Correct Responses Expected by Chance
Q 7. "In the next row, put a mark on the 4."	94	100	20
Q 8. "In the box where the ducks are, put a mark on 56."	74	79	25
Q 9. "See the box where the grapes are. Now write the number 5 in the space beside the grapes."	88	94	0
Q 10. "Find the box where the flag is. Mark the number that comes next after 8 when you are counting."	76	81	25
Q 25. "In the box where the book is, write eighty-one."	62	66	0

seem to reveal the rote knowledge these children had of certain basic social-cultural number symbols (see Table 1).

Table 1 indicates that the children did considerably better on questions 7, 8, 9, 10, and 25 of the MRT, FA, T5N than could be expected by chance. Such responses would tend to indicate that the children in our sample had been socialized to the basic numbers, 4, 5 and 6, 5, 9, and 8 and 1 respectively. This sample of children had been socialized to number symbols for 7½ months prior to the testing. The fomal instrument of this socialization was the kindergarten level workbook of the Greater Cleveland Mathematics Program, 1961. Although Table 1 would indicate that the children did possess a degree of rote knowledge of the number symbols which they were able to write and recognize in print, this knowledge was of little operational value to them: i.e., only eight of the children who were exposed to the structured interview were able to demonstrate that they had reached the stage of attainment of the concept of number. The only rationale which may be given in support of socialization to number symbols is that it provides children with the culturally accepted symbols with which to perform operations after the attainment of the concept of number. For children at the stage of attainment we can include addition and subtraction as goals for a mathematics curriculum.

Throughout the research, the author's concern was with the developmental status of the child and the mathematics curriculum for kindergarten (6:423–34). More precisely, the concern has been with number readiness. The recommendations are based upon the Piagetian qualitative stage-dependent theory within which the author tried to assess the child's cognitive structure at a certain time or development (6:424). The curriculum recommendations for mathematics in the kindergarten are in accord with the cognitive status of the children as demonstrated through evaluation by means of the Piagetian model.

FOOTNOTE

1. Reproduced from Metropolitan Readiness Tests, copyright 1965 by Harcourt, Brace and World, Inc. Reproduced by special permission of the publisher.

REFERENCES

1. Beth, E. W.; Piaget, J., *Mathematical Epistemology and Psychology* (translated from French by W. Mays), D. Reidel Publishing Company, Dordrecht-Holland; Gordon and Breach Science Publishers, New York, 1966.
2. Flavell, J. H., *The Developmental Psychology of Jean Piaget*, Van Nostrand Company, Inc., Princeton, New Jersey, 1963.
3. Hildreth, G. H.; Griffiths, N. L.; McGauvran, M. E., "Metropolitan Readiness Tests, Form A," Harcourt, Brace and World, Inc., New York, 1965.
4. Piaget, J., *The Child's Conception of Number*, Routledge and Kegan Paul Ltd., London, England, 1952.
5. Piaget, J., *Psychology of Intelligence*, Littlefield, Adams and Co., Totowa, New Jersey, 1966.
6. Sigel, I. E.; Hooper, F. H. (eds.), *Logical Thinking in Children*, Holt, Rinehart, and Winston, Inc., New York, 1968.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Field Dependence–Field Independence in 5- to 10-Year-Old Children

GORDON S. BIGELOW, Kent State University

ABSTRACT

A final random sample of 160 Provo, Utah, elementary school children ages 5 to 10, was included in a balanced cell format to investigate the effects of school as reflective of socioeconomic level, age, sex, and intelligence on performance on the Children's Embedded Figures Test (CEFT), a measure of Witkin's field-dependence-independence hypothesis appropriate for use with young children. Results implied that (1) there may be differences on criterion performance according to relative social class; (2) age is the single best predictor of criterion performance, as significant differences proceeding from relative field-dependence to greater field-independence with increasing age were found; (3) there is a similar developmental pattern to increase in verbal intellectual power and increase in relative field-independence when all ages are considered together; (4) there were no significant similarities between field-dependence-independence and verbal intelligence within three of four specific ages investigated; (5) no sex differences in criterion performance were found.

THE RELATIONSHIP between perceptual style and personality characteristics forms the basis of Witkin's (5,6) studies of body-in-field or field-dependence-independence. The Witkin hypothesis infers that the degree to which an individual involves the surrounding field as referants in decision making is directly related to a series of personality characteristics such as dependence, submission to authority, and anxiety (6). In his two major published studies, Witkin (6, 7) employed several tests to differentiate among field-dependence and field-independence, only one of which, the Embedded Figures Test (EFT), required no equipment other than a pencil and a printed booklet. Ease of administration and relatively high correlation with other Witkin tests made the EFT the most widely used research instrument in this area. Due to the high level of perceptual sophistication required, however, the EFT was found to be inappropriately complex for children younger than age 10. Therefore, information related to this dimension of perceptual style either was restricted to people over 10 or was inferred from results of administered projective techniques, a method soundly criticized in the literature (4, 8, 9).

In 1963, an EFT for young children was developed at Down State Medical Center, State University of New York, which lowered the minimum age for objective identification of cognitive style from 10 to 5 (2). The Children's Embedded Figures Test (CEFT) correlated on the order of .70 with the EFT (2) and made it possible to fill in some of the gaps in knowledge of cognitive style with younger children. Few research reports quoting use of this instrument have appeared in the literature. Maccoby (3) used the CEFT in an investigation of activity level and intelligence in young children as did Corah (1) while investigating field-dependence-independence between parents and children.

Problem

The present study was initiated as an attempt to lower the data level from age 10 to age 5 for the Witkin hypothesis. The purpose of this study was to ascertain the degree to which differences in field-dependent and field-independent modes of perceiving, as measured by the CEFT, were related to socioeconomic class, age, sex, and intelligence in 5- to 10-year-old children in two contiguous Provo, Utah, elementary schools. It was hypothesized that, from CEFT results, no significant differences would exist between (1) the two schools, (2) males and females, (3) upper and lower IQ groups, and (4) interactions of age, sex, school, and IQ variables. Also, it was hypothesized that children in the youngest age group would score lower (more field-dependence) than would children in the oldest age group investigated.

Method

All 5—10-year-old students in four age spans were included (60–74 months, 75–89 months, 90–104 months, and 105–119 months). Each group was broken into subgroups by school and then by sex, making a total of 16 subpopulations. From each of the sixteen substrata, twenty children were randomly selected and administered the Pea-

Table 1.—CEFT Performance Between Schools with ISP and with IQ Covariate

Covariate	Mean Square	F ratio	Sig.
ISP	57.1289	3.96	.05
IQ	95.5431	6.75	.05

body Picture Vocabulary Test (PPVT) (N=320). The upper and lower quartile (five children each) of each group were separated according to intelligence test results. The final sample, then, included thirty-two cells of five in each cell broken down into all possible combinations of age (4 strata), sex (2 strata), school (2 strata), and verbal intelligence (2 strata) (4x2x2x2). Each of the final 160 children was administered the CEFT. Further, the school records for each child were researched for information to be included in the Hollingshead Index of Social Position (ISP) if use of this instrument was necessitated by revealed interschool population differences. Hypotheses were tested through use of multiple classification analysis of variance and multiple classification analysis of covariance. Where significant differences were found among strata involving more than two variables (age), orthogonal comparison tests of significance were applied to ascertain degree and location of significance. The .05 level was used as the criterion of significance.

Results

A significant difference between school groups enabled the use of the ISP in an attempt to ascertain whether or not social position differences existed between the schools. Significance was found in social position indicating that students in the school containing children from relatively higher social classes scored more field-independently than did students from the school containing generally lower social class children (according to ISP scores) (see Table 1). School differences in CEFT performance were even more significant when IQ was used as a covariate (see Table 1). Lower social position students showed more field-dependence, and higher social position students showed greater field-independence, a result which was further reinforced when efforts of IQ were held constant.

There were no significant differences between CEFT performance of boys and girls or of high and low IQ groups (see Table 2). A highly significant difference within the age variable reflected greater field-dependence at the earliest age level and relatively greater field-independence at the oldest age level (see Table 2).

Orthogonal comparisons run among the four age groups to ascertain locus of significance are presented in Table 2. When social position and IQ were employed as covariates, significance levels were again discovered and investigated by orthogonal comparisons (see Tables 3 and 4). F ratios and significance levels remained relatively constant comparing results with and without covariates. Highly significant differences in performance on the CEFT between the two younger and two older groups suggested that performance is much more field-independent at older ages and that the increase in performance seems most sharp between pre- and post-8-year-old chil-

Table 2.—Multiple Classification Analysis of Variance for Schools, Age, Sex, and IQ for CEFT Performance by 160 5–10 Year Old Children

Source	df	Mean Square	F ratio
Schools	1	74.2562	5.153*
Sex	1	0.5062	0.009
Sch x Sex	1	56.4062	3.915
Age	3	643.8896	37.825**
Orth. Comp.1[a]	3	1644.8	96.62**
Orth. Comp.2	3	31.25	1.84
Orth. Comp.3	3	255.6	15.01*
Sch x Age	3	17.0229	1.181
Sex x Age	3	2.5062	0.278
Sch x Sex x Age	3	9.0062	0.625
IQ	1	166.0562	3.510
Sch x IQ	1	47.3062	3.283
Sex x IQ	1	3.3062	0.061
Sch x Sex x IQ	1	54.0552	3.751
Age x IQ	3	54.9229	6.167
Sch x Age x IQ	3	8.9062	0.618
Sex x Age x IQ	3	21.4729	19.707*
Sch x Sex x Age x IQ	3	1.0896	0.076
Error	128	14.4094	
TOTAL	159		

[a] Orthogonal Comparison 1 (Age group 1 and 2 to 3 and 4), Orthogonal Comparison 2 (Age Group 1 to 2), Orthogonal Comparison 3 (Age Group 3 to 4).

* Significant at .05 level.

** Significant at .01 level.

dren. Sex × age × IQ interaction proved significant and non-interpretable.

Discussion

Use of the ISP as a covariate enabled discovery of possible differences in global-analytical style between levels of social strata. Students in School A received CEFT scores reflecting a significantly greater degree of field-dependence than the CEFT scores of students at School B. Further, School A students received social position scores significantly lower than did School B students (F=6.789, significant at .01) which ostensibly means that School A students came from a lower socioeconomic level, on the average, than their School B counterparts. In other words, lower social position students, relatively speaking, showed more field-dependence, and higher social position students showed more field-independence. This is a novel result in field-dependent-independent research, and one which bears further investigation. If this dichotomy proves to be replicable, it will carry with it cultural and interclass implications. For example, one might conjecture that high social position children are exposed to more situations fostering development of ability to function independently of environmental support or perhaps this could arise as a result of coming from families with relatively high need for achievement indexes in which pressure to succeed is greater. Perhaps, on the other hand, children from lower social strata are not subjected to as much pressure to achieve. Lower social position children who lived in neighborhoods of relatively lower socioeconomic class families showed more field-dependence. Yet, those children in the more favorable district who were rated to be in the lower socioeconomic classes did not reflect this same field-dependence and, in fact, performed more like the higher social class children in whose neighborhoods they lived. When all children are considered, there is little difference in mean CEFT performance between social classes, and the difference is only apparent when further subdivided by neighborhood (schools) area.

Table 3.—Analysis of Covariance of CEFT Scores of Ss in Each of Four Age Groups With Social Position as Covariate

Source	df	Mean Square	F ratio
Age	3	524.6857	32.70**
Orth. Comp.1	3	1644.8	102.508**
Orth. Comp.2	3	31.25	1.95
Orth. Comp.3	3	255.6	15.93*
Error	2	16.0456	
TOTAL	5		

* Significant at .05 level.
** Significant at .01 level.

Table 4.—Analysis of Covariance of CEFT Scores of Ss in Each of Four Groups With IQ as Covariate

Source	df	Mean Square	F ratio
Age	3	549.9557	22.12*
Orth. Comp. 1	3	1644.8	66.16**
Orth. Comp. 2	3	31.25	1.26
Orth. Comp. 3	3	255.6	10.28*
Error	2	24.8628	
TOTAL	5		

* Significant at .05 level.
** Significant at .01 level.

Highly significant differences in CEFT performance by age groups was expected, with the general trend going toward increasing field-independence as age increased. This result underscored the fact that children proceed perceptually from global to less global styles during the developmental years 5 to 10. Within each age group, however, considerable range of expression was evident and it was obvious that, even as early as age 5 to 6, some children have achieved a degree of field-independence which some other children may never attain. The point of initial divergence of analytical style from the universal perceptual globality of newborns, therefore, seems to take place prior to age 5 according to data produced by this study. However, the increase in average CEFT performance from ages 5–6 to 6–7 was small and not statistically significant, whereas the increase from 7–8 to 8–9 was large and highly significant. Perhaps before ages 5–6 relatively few children develop substantial perceptual style changes whereas the average child's development seems to spurt ahead between ages 7 and 9. Additional research is needed before this generalization can be discussed with confidence.

In this regard, the significant relationship between overall CEFT and PPVT raw scores seemed to indicate that the abilities tested by each are developmental in nature and apparently develop, on the average, within the same space of time. The lack of consistent relationships between CEFT and IQ scores, therefore, did not necessarily mean that there was no relationship between perceptual field style and what is known in our culture as verbal intelligence.

This study underscored the inappropriateness of comparing embedded figures scores which are converted from raw scores in order to correct for the effects of age. In almost every multiple regression analysis by independent variable, age was shown to be the best single predictor of criterion performance. Therefore, to compare these scores with others for which age is controlled is to be criticized. If the abilities are separate, they at least proceed along similar developmental paths between ages 5 and 10. Further, as Witkin (6) points out, embedded figures test performance seems to reach adult levels by age 15 (the age at which intelligence growth is beginning to stabilize as well). On the other hand, the large sex differences noted in EFT performance during adolescence and adulthood seems to indicate that these two variables are at least partially dissimilar in nature.

Also, it must be recognized that the significant correlations between PPVT raw and CEFT scores do not necessarily mean there is a significant relationship between field-dependence-independence and verbal intelligence. CEFT related .64 and PPVT raw related .67 with age, indicating the possibility that age may be producing much of the raw scores. In this vein, inspection of Table 5 is important. Within the three lowest age groups no relationship was revealed between CEFT and PPVT raw scores. Only in age group four was the significance level reached. In other words, in three of the four age groups CEFT performance was independent of performance on the PPVT whether raw or IQ scores were used for comparison. Field-dependence-independence and verbal intelligence proceed along generally similar developmental paths between ages 5 and 10. Gross similarities seem to exist, but within specific ages the similarities do not appear. A justifiable conclusion seems to be that both verbal intelligence and global-analytical cognitive style are traits which mature in individuals at similar rates. When age is not a factor, however (when one looks at performances within a specific age level), cognitive style and verbal intelligence are less likely to be related. This appears to partially replicate former studies with older Ss which find little or no relationship between these two traits.

Table 5.—Correlation Coefficients Between CEFT Scores and PPVT Raw and PPVT IQ Scores for Age Groups

Group	N	PPVT Raw	PPVT IQ
60-74 mos	40	.05425	.04824
75-89 mos	40	.25932	.24479
90-104 mos	40	.16387	.14815
105-119 mos	40	.51940**	.55235**

** Significant at .01 level.

Table 6.—Means of CEFT Scores by Sex and by Age Group (N=20 each group)

Age Group	Male	Female
60-74 mos	7.200	6.500
75-89 mos	8.100	8.100
90-104 mos	11.850	12.350
105-119 mos	15.800	15.550

Sex differences, as found in previous studies, were not produced in this study. In fact, the differences in mean performance for boys and girls never seemed to even commence divergence in the four age groups studied (see Table 6). By age 10, then, boys and girls still seem to use similar cognitive global-analytical modes. The sex differences reported by Witkin (6) occurred most markedly after age 10, and the present study merely lowers the floor in this area.

In summary, then, results implied that (1) there may be differences in criterion performance according to relative social class; (2) age is the best predictor of criterion performance, as significant differences proceeding from relative field-dependence to greater field-independence with increasing age were found; (3) there is a similar developmental pattern to increase in verbal intellectual power and increase in relative field-independence when all ages are considered together; (4) there were no significant similarities between field-dependence-independence and verbal intelligence within three of four specific ages investigated; (5) no sex differences in criterion performance were found.

REFERENCES

1. Corah, N. L., "Differentiation in Children and Their Parents," *Journal of Personality,* 33:300–308, 1965.
2. Karp, S. A.; Konstadt, N. L., *Manual for the Children's Embedded Figures Test,* Cognitive Tests, New York, 1963.
3. Maccoby, E. E.; Dowley, E. M.; Jagen, J. W., Gegerman, R., "Activity Level and Intellectual Functioning in Normal Preschool Children," *Child Development,* 36: 761–770, 1965.
4. Postman, L.; Witkin, H. A., and others, *Personality Through Perception: Book Review, Psychological Bulletin,* 52:79–83, 1955.
5. Witkin, H. A., "Individual Differences in Ease of Perception of Embedded Figures," *Journal of Personality,* 19:1–15, 1950.
6. Witkin, H. A.; Lewis, H. B.; Hertzman, M.; Machover, K.; Meissner, P. B.; Wapner, S., *Personality Through Perception,* Harper, New York, 1954.
7. Witkin, H. A.; Dyke, R. B.; Faterson, H. F.; Goodenough, D. R.; Karp, S. A., *Psychological Differentiation,* Wiley, New York, 1962.
8. Zigler, E., "A Measure in Search of a Theory?" *Contemporary Psychology,* 8:133–135, 1963.
9. Zigler, E., "Zigler Stands Firm," *Contemporary Psychology,* 8:459–461, 1963.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Characteristics of Students Who Make Accurate and Inaccurate Self-Predictions of College Achievement

KARL E. KEEFER, The University of Tennessee

ABSTRACT

College students ($N = 198$) predicted their grades at the beginning of four grading periods of a school year. Average discrepancy scores (ADS) between the self-predicted grade point average (SP GPA) and achieved grade point average (Ach GPA) were obtained for all students. Students in the top quartile (inaccurate predictors) were compared with students in the bottom quartile (accurate predictors) on five dimensions. No differences were found in sex or age. Accurate predictors tended to differ from inaccurate predictors in academic classification ($p < .01$), academic achievement ($p < .001$), and self-concept ($p < .001$).

THE PREDICTION of academic performance is a *leitmotif* of the educational world, and nowhere more so than in the college and university. Admissions officers, registrars, personnel deans, counselors, teachers, and others involved with students utilize various instruments ranging from statistical projections to intuitive feelings in an effort to peer into the academic future of this or that student, or of a group of students. These efforts have met with varying degrees of success (5, 11, 14).

Students also predict their own achievement, usually on an informal basis, and often without awareness of the thought processes involved. When self-prediction is formalized and the student is made aware of what he is doing, the resultant data are of the sort utilized in the level of aspiration studies which were so prominent a few decades ago (3, 4, 6, 7, 12, 13). These and more recent investigations (1, 2, 8, 9, 10) suggest that some systematic differences may be found between students who make accurate and those who make inaccurate predictions of their own achievement.

The present study is an analysis of the self-predictions of a group of college students to see whether differences between accurate and inaccurate predictors appeared in certain readily observable characteristics. The dimensions which were chosen for analysis, based in part upon the results of level of aspiration studies and in part upon the availability of data, were the following: (1) sex; (2) age; (3) academic classification; (4) academic achievement; (5) self-concept. Data for the first four were obtained from the college records.

Academic achievement was identified by three measures: (1) scores on the American College Test (ACT), taken by most students prior to admission; (2) the total high school GPA (HS GPA); (3) the mean achieved college GPA during the year of the investigation (Ach GPA). All GPA's were computed on the basis of a 4-point scale ranging from $A = 4$ to $F = 0$.

Six indicators of self-concept were utilized: (1) the self-predicted GPA (SP GPA); (2) self-reports of feelings of academic success at the close of the investigation; (3) self-reports of feelings of predictive accuracy at the close of the investigation; (4) an operational measure of academic success during the investigation; (5) an operational measure of predictive accuracy during the investigation; (6) scores on Bills' Index of Adjustment and Values.

Procedure

Subjects for the study were 198 middle class liberal arts undergraduate volunteers, who were asked at the beginning of the school year to complete a card on which they listed all of the courses for which they were currently registered and the grade they expected to receive in each course at the end of the first grading period. This procedure was repeated at the beginning of each of the other three grading periods of the 2-semester school year. In recognition of the possibly threatening nature of this request, provision was made for the coding of all responses, and students were assured of anonymity throughout the investigation.

For each student, an average discrepancy score (ADS) was obtained by computing the mean of the differences between the SP GPA and the Ach

Table 1.—Sex, Age, and Academic Classification of Accurate and Inaccurate Predictors

Variable	Quartile Upper	Quartile Lower	Chi-square
SEX			
Male	32	23	
Female	18	27	3.386
AGE			
Under 18	6	3	
18	15	13	
19	9	7	
20	9	10	
21	1	15	
Over 21	10	12	4.292
ACADEMIC CLASSIFICATION			
Freshman	28	15	
Sophomore	10	7	
Junior	8	15	
Senior	4	13	11.354*

* $p < .01$

GPA for the four grading periods (mid-term and final predictions and grades for two semesters). The group was then ranked from the largest to the smallest ADS, and the top and bottom quartiles were chosen for further analysis. The mean ADS for the top quartile was .68, and for the bottom quartile .20. The difference was significant at the .01 level. The standard deviations for the two groups were .22 and .04 respectively, a difference which yielded a significant F ($p<.01$) indicating that the less accurate predictors were more variable in the accuracy of their predictions than the more accurate predictors.

Each of the Ss in the top and bottom quartiles was identified on each of the dimensions and measures previously listed insofar as data were available. Differences were computed and tested for significance by either the t-test (for measurable characteristics) or the chi-square test (for classifications). The results were interpreted on the basis of the null hypothesis of no significant difference between the upper and lower quartiles.

Findings

The data with regard to sex, age, and academic classification are presented in Table 1. There were no significant differences in sex or age between students who were accurate and those who were inaccurate in their self-predictions. In academic classification, a significant difference was found, and the distribution indicates that freshmen tend to be more inaccurate in their self-predictions (upper quartile of the ADS continuum) and that juniors and seniors tend to be more accurate (lower quartile).

The data on achievement appear in Table 2. It seems clear that, by all three measures, the students who are accurate predictors of their own achievement tend to achieve at a significantly higher level than those who are inaccurate. This difference is reflected by a 3-point differential in

Table 2.—Academic Achievement of Accurate and Inaccurate Predictors

Variable	Upper N	Upper Mean	Upper SD	Lower N	Lower Mean	Lower SD	t	p
ACT	40	16.80	4.30	36	20.10	4.01	3.403	$< .01$
HS GPA	49	2.30	.55	49	2.82	.67	4.317	$< .001$
Ach GPA	50	2.00	.53	50	2.81	.43	8.371	$< .001$

Table 3.—Self-Concept of Accurate and Inaccurate Predictors

Variable	Quartile						t	p
	Upper			Lower				
	N	Mean	SD	N	Mean	SD		
SP GPA	50	2.59	.30	50	2.84	.37	3.794	< .001
IAV II	45	164.6	22.4	50	171.2	21.4	1.452	NS
IAV IV	45	49.2	14.4	50	39.1	14.6	3.355	< .01

ACT scores, half a letter in high school grades, and almost a whole letter in college grades.

The various measures of self-concept appear in Tables 3 and 4. The SP GPA itself (Table 3) shows a significant difference between students with a large ADS and those with a small ADS. The accurate predictors tend to predict higher achievement for themselves than do the inaccurate predictors, a judgment which conforms to the reality of the situation, and which may indicate a greater degree of self-confidence which enables them to be more realistic in their predictions.

In Table 3 evidence appears of greater self-acceptance on the part of students making accurate predictions, as measured by Bills' Index of Adjustment and Values. This instrument yields four scales, two of which are indicative of one's feelings about himself. Scale II is a direct reflection of self-acceptance through a range of five responses ranging from "I very much dislike being as I am" to "I like very much being as I am," made to a series of forty-nine descriptive adjectives. On this scale, although there was a tendency for inaccurate predictors to verbalize lower self-acceptance than accurate predictors, the difference was not significant.

Scale IV measures the self-acceptance variable indirectly. It is derived by taking the difference between Scale I, which pictures the S's actual self-concept, and Scale III, which depicts his ideal self-concept. This difference may better reflect the individual's true, though often unverbalized, feelings about himself than does the direct approach of Scale II, where he may hesitate to express strong like or dislike of some of his characteristics. In interpreting the scores obtained from this scale, it is important to note they are the inverse of Scale II; that is, larger values indicate lower self-acceptance and smaller ones indicate greater self-acceptance. Although the groups in this study did not differ significantly on Scale II, on Scale IV the more accurate predictors were significantly more self-acceptant. (The Pearson r for Scales II and IV for the entire N in the study was —.67, indicating a moderate inverse relationship between these measures.)

In Table 4 differences are presented between accurate and inaccurate self-predictors in the self-concept as inferred from self-reports of feelings of academic success or lack of success and of predictive accuracy or inaccuracy at the end of the experiment. It is assumed that such feelings tend to reflect a positive or negative self-concept. Students with a lower ADS (the accurate group) perceived themselves as significantly more successful and accurate than students with a higher ADS (the inaccurate group).

The validity of these self-reports was confirmed when success was defined operationally as an Ach GPA equal to or greater than the SP GPA, and accuracy was defined operationally as an Ach GPA within one-fourth of a grade point above or below the SP GPA. Students with a lower ADS were operationally successful and accurate a significantly larger number of grading periods than the higher ADS group.

Discussion

These data indicate that while college students who make accurate predictions of their own academic achievement do not differ in sex or age from those whose predictions are inaccurate they do differ in other easily measurable ways. They tend to be:

1. *upperclassmen rather than freshmen;*
2. *higher in academic achievement, as measured by ACT test scores, high school GPA, and college GPA;*

3. *more consistently accurate in their predictions, and to predict higher GPA's;*
4. *more self-acceptant, as inferred from Bills' Index of Adjustment and Values and from their self-reports of feelings of academic success and predictive accuracy;*
5. *operationally more successful in achieving their stated goals and operationally more accurate in predicting their achievement.*

Table 4.—Success and Accuracy of Accurate and Inaccurate Predictors

Variable	Quartile Upper	Quartile Lower	Chi-square
SELF-REPORTED SUCCESS			
Successful	23	43	19.062*
Unsuccessful	22	7	
No report	5	0	
SELF-REPORTED ACCURACY			
Accurate	17	45	33.676*
Inaccurate	28	5	
No report	5	0	
OPERATIONAL SUCCESS (number of grading periods)			
0	33	4	38.318*
1	10	17	
2	3	13	
3	2	12	
4	2	4	
OPERATIONAL ACCURACY (number of grading periods)			
0	27	1	72.184*
1	17	2	
2	6	14	
3	0	27	
4	0	6	

$^*p < .001$

The scope of this study was too limited to permit widespread generalization from these results. Nevertheless, the size of the differences on many of the variables suggests that those assisting college students in academic planning might do well to give serious consideration to student self-predictions when they are made by upperclassmen, high achievers, self-acceptant students, or students who have shown themselves to be successful in achieving their goals and accurate in predicting their achievement. The self-predictions of the student combining several of these characteristics may be especially worthy of such consideration.

REFERENCES

1. Borislow, B., "Self-Evaluation and Academic Achievement," *Journal of Counseling Psychology,* 9:246–254, 1962.
2. Byers, J. L., "Study of the Level of Aspiration of Academically Successful and Unsuccessful High School Students," *California Journal of Educational Research,* 13:209–216, 1962.
3. Frank, J. D., "Recent Studies of the Level of Aspiration," *Psychological Bulletin,* 38: 228–235, 1941.
4. Gardner, J. W., "The Relation of Certain Personality Variables to Level of Aspiration," *Journal of Psychology,* 9:191–206, 1940.
5. Giusti, J. P., "High School Average as a Predictor of College Success: A Survey of the Literature," *College and University,* 39:200–209, 1964.
6. Hilgard, E. R., "Success in Relation to Level of Aspiration," *School and Society,* 55:423–428, 1942.
7. Lewin, K.; Dembo, T.; Festinger, L.; Sears, P. S., "Level of Aspiration," in Hunt, J. McV. (ed.), *Personality and the Behavior Disorders,* Ronald Press, New York, 1944, pp. 333–378.
8. Locke, E. A., "Relationship of Intentions to Level of Performance," *Journal of Applied Psychology,* 50: 60–66, 1966.
9. Payne, D. A., "Concurrent and Predictive Validity of an Objective Measure of Academic Self-Concept," *Educational and Psychological Measurement,* 22:773–780, 1962.
10. Red, S. B.; McCary, J. L.; Johnson, B., "Study of the Relationship Between Aspiration Levels and Academic Achievement," *Journal of Educational Research,* 55:159–163, 1961.
11. Richards, J. M., Jr.; Holland, J. L.; Lutz, S. W.; "Prediction of Student Accomplishment in College," *Journal of Educational Psychology,* 58: 343–355, 1967.
12. Rotter, J. B., "'Level of Aspiration as a Method in the Study of Personality: I. A Critical Review of Methodology," *Psychological Review,* 49:463–474, 1942.
13. Rotter, J. B., "Level of Aspiration as a Method in the Study of Personality: II. Development and Evaluation of a Controlled Method," *Journal of Experimental Psychology,* 31:410–422, 1942.
14. Yonge, G. D., "Students: Prediction," *Review of Educational Research,* 35:255–256, 1965.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Faculty Allocation of Effort and Student Course Evaluations

CLAUDE W. GRANT, University of Utah

ABSTRACT

Two types of data were obtained on teaching faculty at the University of Utah, student evaluations of courses, and faculty allocations of time distributed among academic activties. The data were subjected to treatment by analysis of variance procedures. It was found that faculty rank was not related to student evaluations of courses on any of the five course evaluation scales and in general there was little relationship between faculty allocations of time and student evaluations of courses. It was noted, however, that students in courses with large enrollments considered faculty more "prepared" than students in courses with smaller enrollments. There was some evidence that as faculty time allocated to research and writing increased, student ratings of courses decreased.

THIS PRESENTATION is a discussion of the interrelationship of two types of data obtained about teaching faculty at the University of Utah, autumn quarter, 1968. The two types of data are: 1. student evaluations of courses (1) and 2. faculty allocations of time distributed among university academic activities (2).

The existence of these data allowed the following question to be asked: *"Is there a relationship between the activities in which faculty distribute their faculty time and the evaluations made of faculty members' courses by students?"* For purposes of this study the question was broken down into the following sub-questions:

1. *Is the rank of the instructor a significant variable in student evaluations of a course?*

2a. *Is the level of the course (lower, upper, or graduate division) a significant variable in student evaluations of the course?*
  b. *Is there a significant interaction between the level of the course and the rank of the instructor teaching the course?*

3a. *Is the percent of an instructor's total time spent in classroom instruction (including course preparation and advising students within courses) a significant variable in student evaluations of his courses?*
  b. *Is there a significant interaction between this time factor and the instructor's rank?*

4a. *Regarding specific courses, is the time allocated to a specific course by an instructor a significant variable in student evaluations of that course?*
  b. *Is there a significant interaction between this time factor and the instructor's rank?*

5a. *Is number of students enrolled in a course a significant factor in student evaluations of that course?*
  b. *Is there a significant interaction between the enrollment in the course and instructor's rank?*

6a. *Is the amount of time instructors spend in supporting activities such as program supervision, counseling students, committee work, a significant variable in student evaluations of courses?*
  b. *Is there a significant interaction between time spent in supporting activities and instructor's rank?*

7a. *Is time spent in non-sponsored (nonfunded) research and writing a significant variable in student evaluations of courses?*
  b. *Is there a significant interaction between this time factor and an instructor's rank?*

8a. *Is time spent in sponsored (funded) research and writing a significant variable in student evaluations of courses?*
  b. *Is there a significant interaction between this time factor and an instructor's rank?*

Procedure

Course Evaluation. The course evaluation form was developed following the model of Osgood's Semantic Differential Scales (3). There were seven scales in the evaluation form: "Preparation," "Explanation," "Responsiveness," "Stimulation," "Examinations," "Text," and Recomamended." In this study the scales on "Examinations" and "Text" were excluded.

Scales were printed on IBM data card-size forms identified by the instructor's name, course, and credit value. Evaluation cards were sent to the students in their pre-registration packets during the latter part of the autumn quarter. They were requested to complete the forms and return them with their pre-registration materials. The evaluations were anonymous.

Only those courses in which 75 percent of the enrolled students completed course evaluation forms were included in this study. Graduate students did not respond as cooperatively as undergraduate students, and as a result, graduate courses are not represented to the same extent as undergraduate courses are.

Faculty Effort. The form on which faculty were asked to allocate their time was programmed to reduce to a minimum the time required to complete the form. The form was printed on an 8 x

Table 1.—Faculty Effort and Course Evaluation

COURSE EVALUATION SCALE→ FACULTY ACTIVITY	RECOMMENDED	EXPLANATION	RESPONSIVENESS	STIMULATION	PREPARATION
RANK	N.S.	N.S.	N.S.	N.S.	N.S.
RANK BY COURSE LEVEL					
A. Rank	N.S.	N.S.	N.S.	N.S.	N.S.
B. Course Level	N.S.	N.S.	N.S.	N.S.	N.S.
C. Interaction A x B	N.S.	N.S.	N.S.	N.S.	N.S.
RANK BY % OF TIME IN CLASSROOM INSTRUCTION					
A. Rank	N.S.	N.S.	N.S.	$P < .05$	N.S.
B. Percent of Time	N.S.	$P < .05$	N.S.	$P < .01$	$P < .05$
C. Interaction A x B	N.S.	N.S.	N.S.	N.S.	$P < .05$
RANK BY % OF TIME IN SPECIFIC COURSES					
A. Rank	$P < .05$	N.S.	N.S.	N.S.	N.S.
B. Percent of Time	N.S.	N.S.	N.S.	N.S.	N.S.
C. Interaction A x B	N.S.	N.S.	N.S.	N.S.	N.S.
RANK AND CLASS SIZE					
A. Rank	N.S.	N.S.	N.S.	N.S.	N.S.
B. Class Size	N.S.	N.S.	N.S.	N.S.	$P < .01$
C. Interaction A x B	N.S.	N.S.	N.S.	N.S.	N.S.
RANK AND ACTIVITIES SUPPORTING INSTRUCTION					
A. Rank	N.S.	N.S.	N.S.	N.S.	N.S.
B. Activities	N.S.	N.S.	N.S.	N.S.	N.S.
C. Interaction A x B	N.S.	N.S.	N.S.	N.S.	N.S.
RANK AND NON-SPONSORED RESEARCH AND WRITING					
A. Rank	N.S.	N.S.	N.S.	N.S.	N.S.
B. Non-Sponsored Research and Writing	$P < .01$	$P < .01$	$P < .01$	$P < .01$	$P < .05$
C. Interaction A x B	$P < .01$	$P < .06$	$P < .01$	N.S.	N.S.
RANK AND SPONSORED RESEARCH					
A. Rank	N.S.	N.S.	N.S.	N.S.	N.S.
B. Sponsored Research	N.S.	N.S.	N.S.	N.S.	N.S.
C. Interaction A x B	$P < .05$	$P < .05$	$P < .06$	$P < .05$	N.S.

11-inch sheet. Programmed onto the form was the faculty member's name, social security number,

rank, department(s), source(s) of salary, percent of salary from each source, and the courses he was teaching during the autumn quarter with course credit value and number of students indicated.

Faculty members were asked to allocate their time on a percentage basis among the following categories.

1. *Each course taught (including preparation and student advising)*
2. *Graduate student supervision*
3. *Activities supporting instruction (program supervision, counseling, committee work)*
4. *Non-sponsored (nonfunded) research and writing*
5. *Sponsored (funded) research and writing*
6. *University and college administration (excluded in this study).*

Sample. The sample consisted of 685 faculty members, on whom data were available from the faculty effort study and the course evaluation study. By rank, there were 154 professors, 110 associate professors, 131 assistant professors, sixty-three instructors and lecturers, and 225 teaching assistants. The sample size by category varies for each of the questions considered in this presentation because in some instances one or more students did not respond to a specific scale on the evaluation form making it necessary to drop that faculty member from the analysis of data related to that scale.

Two sets of IBM data cards were punched for each faculty member, and were identified for merging purposes by social security number. A one-way analysis of variance program was utilized to examine rank as a variable in student ratings; for each of the other analyses, a two-way analysis of variance program was used. The data were run on a UNIVAC 1108 computer.

Findings. A summary of the findings are presented in Table 1.[1] It is possible, from looking at this table, to determine those instances where variables proved to be significant factors in student evaluations of courses; also, it is possible to observe where significant interaction occurs.[2] Presentation of the findings is provided as responses to the eight questions phrased in the introductory section of this presentation, as follows:

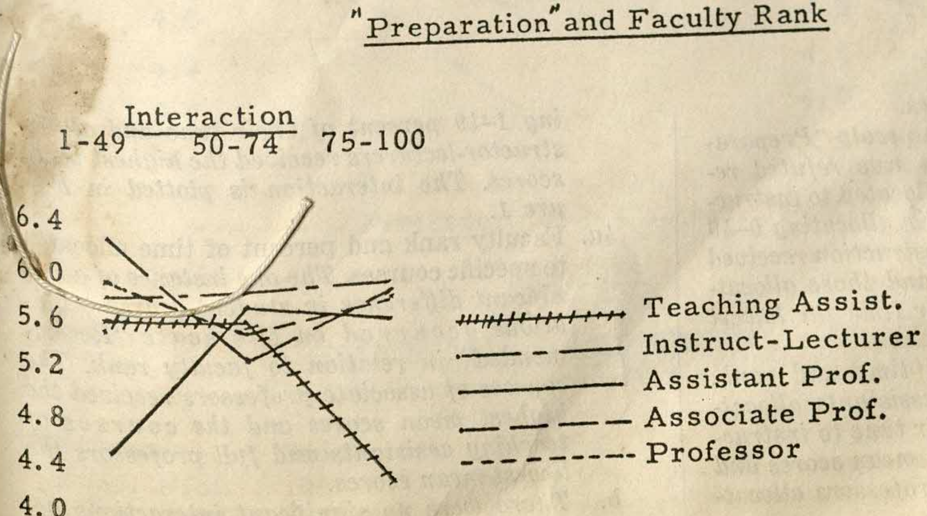
1. Faculty rank. *Observably, from Table 1, the statistical tests used did not upset the null hypothesis that rank of instructor was not a significant factor in student evaluation of courses.*

2a–b. Faculty rank and level of course taught. *Student evaluations of courses were not different by course level, nor were there any significant interactions between instructor's rank and level of course taught.*

3a. Faculty rank and the percent of time allocated to instruction. *In the instance of the scale "Explanation," the null hypothesis regarding percent of time allocated to instruction was not supported. Courses taught by instructors allocating 0–49 percent of their time to classroom instruction received the highest mean scores on this scale. Courses receiving the lowest mean scores were taught by instructors allocating 50–74 percent of their time to classroom instruction.*

In the instance of the scale "Stimulation," the null hypothesis was not supported regarding either faculty rank or percent of time allocated to instruction. Courses taught by teaching assistants received the lowest mean score on this scale and courses taught by instructors and lecturers received the highest mean score. (The instructor-lecturer category is not a pure category. Many faculty in this classification are at the University on a temporary or part-time basis. The rank cannot be considered an indication of professional status in the same sense as other ranks.) Courses taught by faculty allocating 50–74 percent of their time to instruction received the lowest mean "Stimulation" scores and those allocating 75–99 percent of their time received the highest

Figure 1. Percent of Time Allocated to Classroom Instruction and Faculty Rank

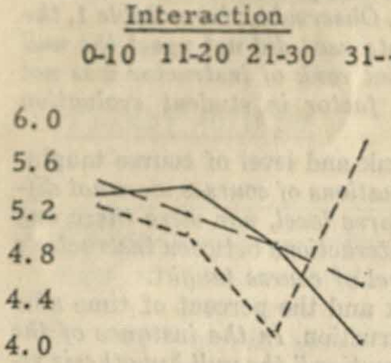


Cell Size

	1-49	50-74	75-100
Teaching Assist.	30	196	2
Instruct-Lecturer	17	20	25
Assistant Prof.	10	46	73
Associate Prof.	10	38	59
Professor	25	43	82

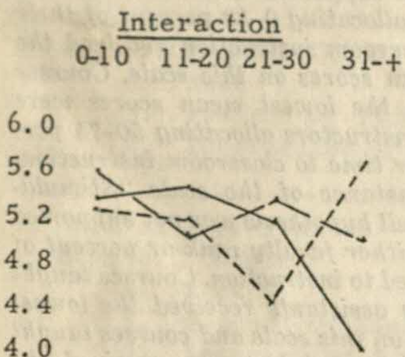
Figure 2. Percent of Time Allocated to Non-Sponsored Research and Writing and Faculty Rank

"Explanation" and Faculty Rank



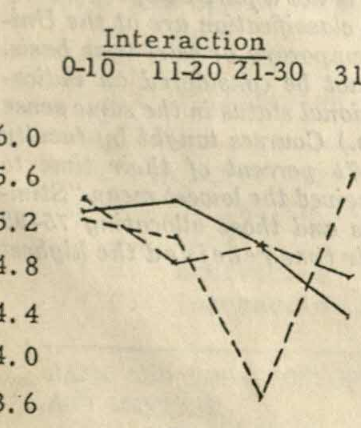
	Cell Size			
	0-10	11-20	21-30	31-+
--------Professor	120	16	9	5
— — —Assoc. Prof.	78	19	9	3
———— Assist. Prof.	78	22	14	15

"Responsiveness" and Faculty Rank



	Cell Size			
	0-10	11-20	21-30	31-+
--------Professor	110	15	9	5
— — —Assoc. Prof.	78	18	9	3
————Assist. Prof.	79	22	14	13

"Recommended" and Faculty Rank



	Cell Size			
	0-10	11-20	21-30	31-+
------- Professor	124	16	9	5
— — — Assoc. Prof.	79	19	9	3
———— Assist. Prof.	80	22	14	15

mean "Stimulation" scores.

In the instance of the scale "Preparation," the null hypothesis was refuted regarding percent of time allocated to instruction. The courses of faculty allocating 0–49 percent of their time to instruction received the highest mean scores and those allocating 50–74 percent of their time the lowest mean scores.

b. *Regarding interaction of time and rank, the courses of teaching assistants allocating 75–100 percent of their time to instruction were given the lowest mean scores and the courses of associate professors allocating 1–49 percent of their time and all instructor-lecturers received the highest mean scores. The interaction is plotted in Figure 1.*

4a. Faculty rank and percent of time allocated to specific courses. *The one instance of a significant difference in student course evaluations occurred on the scale "Recommended" in relation to faculty rank. The courses of associate professors received the highest mean scores and the courses of teaching assistants and full professors the lowest mean scores.*

b. *There were no significant interactions be-*

tween faculty rank and percent of time spent in specific courses on any of the scales.

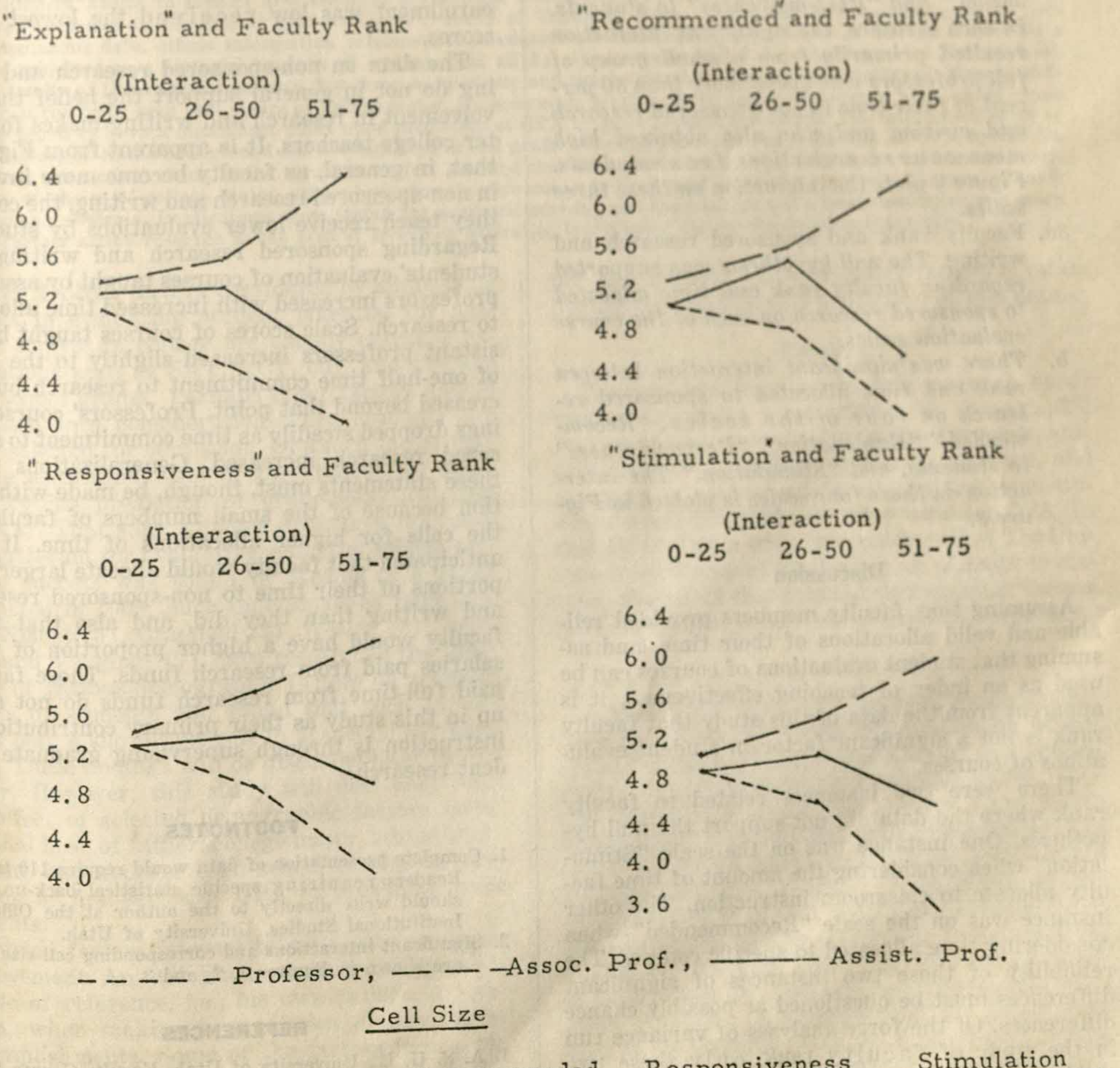
5a. Faculty rank and class size. *The null hypothesis was supported in each analysis except in the instance of class size on the scale "Preparation." The highest mean score on this scale was given to courses in which the enrollment exceeded eighty and the lowest mean score was given to courses with enrollments between ten and twenty-five. Courses with enrollments below ten were excluded from the study.*

b. *There were no significant interactions between rank and class size.*

6a–b. Faculty rank and activities supporting instruction. *The null hypothesis was supported in each instance. Neither rank nor time allocated to activities supporting instruction proved significant variables in student evaluations of courses, nor were there any significant interactions between these two variables.*

7. Faculty rank and non-sponsored research and writing. *Time allocated to non-spon-*

Figure 3. Percent of Time Allocated to Sponsored Research and Writing and Faculty Rank



Cell Size

	Explanation			Recommended			Responsiveness			Stimulation		
	0-25	26-50	51-75	0-25	26-50	51-75	0-25	26-50	51-75	0-25	26-50	51-75
Professor	144	4	2	148	4	2	133	4	2	146	4	2
Assoc. Prof.	96	9	4	97	9	4	95	9	4	95	9	4
Assist. Prof.	116	11	2	118	11	2	115	11	2	117	11	2

sored research and writing proved to be a significant variable in student ratings of courses on each of the five course evaluation scales included in this study. On each scale and for each faculty rank classification, with the exception of a small group of full professors who spent more than 30 percent of their time in non-sponsored research and writing, as the faculty members' time allocated to non-sponsored research and writing increased, the courses they taught received lower mean scale scores.

b. *On three scales, there was a significant interaction between rank and time allocated, these scales were "Recommended," "Explanation," and "Responsiveness" to students. In each instance, the significant interaction resulted primarily from a small group of full professors who spend more than 30 percent of their time in non-sponsored research and writing and who also obtained high mean course evaluations from students. Figure 2 plots the interaction on these three scales.*

8a. Faculty rank and sponsored research and writing. *The null hypothesis was supported regarding faculty rank and time allocated to sponsored research on each of the course evaluation scales.*

b. *There was significant interaction between rank and time allocated to sponsored research on four of the scales, "Recommended," "Explanation," "Responsiveness" to students, and "Stimulation." The interaction on these four scales is plotted in Figure 3.*

Discussion

Assuming that faculty members provided reliable and valid allocations of their time, and assuming that student evaluations of courses can be used as an index of teaching effectiveness, it is apparent from the data of this study that faculty rank is not a significant factor in student evaluations of courses.

There were two instances related to faculty rank where the data did not support the null hypothesis. One instance was on the scale "Stimulation" when considering the amount of time faculty allocate to classroom instruction. The other instance was on the scale "Recommended" when considering time allocated to specific courses. The reliability of these two instances of significant differences must be questioned as possibly chance differences. Of the forty analyses of variance run in the study of faculty rank, only these two reached a level of significance at the .05 probability level. Two significant differences out of forty could occur on a chance basis.

The percent of time allocated to instruction proved to be a significant variable in student evaluations of courses on three scales, "Explanation," "Stimulation," and "Preparation," but not on a direct ratio basis; that is, scale scores did not increase directly with increase in time allocated to instruction. Students rated courses of instructors, who fell in the lowest time category, highest on the "Explanation"and "Preparation" scales. Students rated courses of instructors, who fell in the highest time category, highest on the "Stimulation" scale. The reason for this inconsistency is not at all clear.

It is evident from the data on class size that the size of the class is a factor in student evaluations of the preparedness of the instructor. Courses in which enrollment was high received the highest mean scores, and courses in which enrollment was low received the lowest mean scores.

The data on non-sponsored research and writing do not in general support the belief that involvement in research and writing makes for better college teachers. It is apparent from Figure 2 that, in general, as faculty become more involved in non-sponsored research and writing, the courses they teach receive lower evaluations by students. Regarding sponsored research and writing, the students' evaluation of courses taught by associate professors increased with increased time allocated to research. Scale scores of courses taught by assistant professors increased slightly to the point of one-half time commitment to research but decreased beyond that point. Professors' course ratings dropped steadily as time commitment to sponsored research increased. Generalizations from these statements must, though, be made with caution because of the small numbers of faculty in the cells for higher allocations of time. It was anticipated that faculty would allocate larger proportions of their time to non-sponsored research and writing than they did, and also that more faculty would have a higher proportion of their salaries paid from research funds. Those faculty paid full-time from research funds do not show up in this study as their primary contribution to instruction is through supervising graduate student research.

FOOTNOTES

1. Complete presentation of data would require 110 tables. Readers requiring specific statistical back-up data should write directly to the author at the Office of Institutional Studies, University of Utah.
2. Significant interactions and corresponding cell size data are shown in Figures 1, 2, and 3.

REFERENCES

1. A. S. U. U., University of Utah, *Faculty-Course Evaluation Survey Report, Fall Quarter, 1968–69,* Desert News Press, 1969.
2. Grant, C. W., "Faculty Effort Study," Office of Institutional Studies, University of Utah, Salt Lake City, Autumn Quarter, 1968.
3. Osgood, C. E.; Suci, G. J.; Tannenbaum, P. H., *The Measurement of Meaning,* University of Illinois Press, Urbana, 1957.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

The Effect of Selected Demographic Factors on Prediction of GPA by Male Freshman Students

RICHARD F. SEBOK
Stanislaus State College

ABSTRACT

Student ability to predict first semester grade-point averages (GPA's) was examined in a study of 1,025 male college freshmen in the fall of 1968 at The University of Texas, El Paso. A questionnaire was designed to yield socioeconomic data, ethnic information, information on educational background, and choice of college major. Data were subjected to a factorial analysis of variance and all hypotheses were tested for significance at the .05 level. The results of this analysis indicate that three main effects and two interaction effects were significant. More specifically, it can be be seen that choice of college major, socioeconomic level, and the factor of average high school grade were found to have a significant effect upon prediction of GPA by male freshman students. The interaction effect of the factors of ethnic background and average high school grade were significant in the prediction of GPA by the Ss. A final interaction which had a significant effect on predicted GPA was that of the four main independent variables of choice of college major, ethnic background, socioeconomic level, and average high school grade of male freshman students. Implications were stated for academic advising and financial aid and a recommendation was made for further action research in the area of freshmen aspiration levels using the results found in this study.

THE MAJOR factors considered by admissions officials in evaluating student potential for academic success in college are usually entrance examinations such as SAT, ACT, etc., and average high school grades. Perhaps other factors should be taken into consideration when students are being selected for admission. One such factor may be the aspiration level of the student as measured by his prediction of grade-point averages (GPA's). That such a factor does have relevance as a criterion for evaluating potential college academic success was found by Keefer in a study conducted in 1966. Such individuals as Bernstein (2), Nisbet (5), Bradfield (3), Wesley (8), Astin (1), and Vaughan (7) have researched the influences of various demographic factors on achieved GPA's of students and found that such factors often play a significant role in academic performance. These findings will be discussed later in this paper. However, this study will deal only with the effect of selected demographic factors (educational level of father, college major, ethnic background, and average high school grade) on prediction of first-semester GPA by freshman male students.

Myriad factors interact to determine academic achievement. An individual usually relies upon his frame of reference, i.e., his own particular situation, when making a prediction of his future accomplishments. Some of the factors that enter into his prediction are family circumstance, present socioeconomic conditions, and past experiences. The weight attached by the student to various factors will do much to "color" a prediction. All factors of an individual's conscious awareness interact to form a basis on which he can predicate a prediction. Naturally, self-understanding and personal honesty will be deciding factors in establishing the validity of a prediction of this nature.

Problem

The problem involved in this research paper was to determine what effect such demographic factors as parents' educational backgrounds, academic major in college, ethnic background, and average high school grade have on freshman students' predictions of their first-semester GPA. In the fall of 1968 a study was conducted at The University of Texas at El Paso in an attempt to discover the effect that selected demographic factors have on student prediction of academic success.

Method

A questionnaire designed to yield socioeconomic data, ethnic information, information on educational background, and college major was administered to 1,025 entering freshman males. This questionnaire was the only source from which information was derived; data presented in this paper are based solely on students' responses to the questionnaire. Students were asked on the questionnaire to predict their first-semester GPA. Data were compiled by comparing mean predicted GPA's for students from different backgrounds as determined by these selected independent variables. Analysis of variance and the t-test were used as the statistical tools to obtain results of the effect of the selected demographic factors on predicted GPA.

The following research hypotheses were formulated in null form in order to facilitate statistical treatment of the data:

1. *There will be no significant effect of college*

major of male freshman students on prediction of GPA.

2. *There will be no significant effect of ethnic background of male freshman students on predicted GPA.*
3. *There will be no significant effect of socioeconomic level of male freshman students on predicted GPA.*
4. *There will be no significant effect of average high school grade of male freshman students on predicted GPA.*
5. *There will be no significant interaction effect of college major and average high school grade of male freshman students on prediction of GPA.*
6. *There will be no significant interaction effect of ethnic background and average high school grade of male freshman students on predicted GPA.*
7. *There will be no significant interaction effect of socioeconomic level and average high school grade of male freshman students on predicted GPA.*
8. *There will be no significant interaction effect of college major and ethnic background of male freshman students on predicted GPA.*
9. *There will be no significant interaction effect of college major and socioeconomic level of male freshman students on predicted GPA.*
10. *There will be no significant interaction effect of ethnic background and socioeconomic level of male freshman students on predicted GPA.*
11. *There will be no significant interaction effect of college major, ethnic background, socioeconomic level, and average high school*

Table 1.—Analysis of Variance Showing Significant Effects of Independent Variables of Male Freshman Students on Predicted GPA

Source of Variance N=1,025	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
(CM) College Major	6.23	4	1.52	7.40**
(E) Ethnicity	.46	2	.23	1.09
(SE) Socioeconomic Level	3.87	3	1.29	6.14**
(G) High School Grade-Point Average	18.97	3	6.32	30.07**
CM x E	2.74	8	.34	1.65
CM x SE	.74	12	.06	.30
CM x G	.10	12	.00	.00
E x SE	.62	6	.10	.49
E x G	4.08	6	.68	3.53**
SE x G	3.15	9	.35	1.83
CM x E x SE x G	24.19	72	.34	1.60**
Within Groups	165.08	785	.21	

**$p < .01$

grade of male freshman students on predicted GPA.

Results

The data were subjected to a factorial analysis of variance and all hypotheses were tested for significance at the .05 level. The results of this analysis are presented in Table 1 and indicate that three main effects and two interaction effects were significant. More specifically, it can be seen that choice of college major, socioeconomic level, and the factor of average high school grade were found to have a significant effect on prediction of GPA by male freshman students.

The factors of ethnic background and average high school grade were significant in the prediction of GPA by the Ss. However, the significance of this effect must be attributed to average high school grade since it was found that ethnicity had no significant effect. The final interaction to be considered for effect on predicted GPA was that of the four main factors of choice of college major, ethnic background, socioeconomic level, and average high school grade of male freshman students. This interaction does have a significant effect on prediction of GPA.

The academic fields considered were business, education, engineering, liberal arts, and science. Educational level of father was used as the criterion for determining the socioeconomic level of the Ss. Educational levels considered were 0–8 years of school, some high school or high school graduate, some college or college graduate, and graduate work. Average high school grades considered in the study were A, B, C, and D (see Table 1).

In order to study the effects of various choices of major academic field, the data were also subjected to the t-test. Significant differences in GPA predictions were revealed when predictions were compared on the basis of different academic majors. Engineering majors predicted higher GPA's than did students majoring in business. Table 2 presents comparisons of GPA predictions of engineering majors with those of students majoring in business.

Table 2.—t-Test Showing Significant Difference Between Predicted GPA's of Male Freshman Students Majoring in Engineering and of Those Majoring in Business

GROUP	MEAN	SD	NDF	t
Business	2.44	.45	468	2.21*
Engineering	2.54	.49		

* $p < .05$

Table 3.—t-Test Showing Significant Differences Betwee Predicted GPA's of Male Freshman Students Majoring i Science and of Those Majoring in Business, Education Engineering, and Liberal Arts

GROUP	MEAN	SD	NDF	t
Business	2.44	.45	436	5.17**
Science	2.67	.49		
Education	2.47	.39	299	3.77**
Science	2.67	.49		
Engineering	2.54	.49	388	2.79**
Science	2.67	.49		
Liberal Arts	2.53	.44	431	3.26**
Science	2.67	.49		

** $p < .01$

Science majors predicted higher GPA's than did business majors, education majors, engineering majors, and liberal arts majors (see Table 3). Liberal arts majors predicted higher GPA's than did students majoring in business (see Table 4). Other than the above area of significance, liberal arts majors did not predict significantly higher GPA's than did students in any other major field.

Table 4.—t-Test Showing Significant Difference Between Predicted GPA's of Students Majoring in Liberal Arts and of Those Majoring in Business

GROUP	MEAN	SD	NDF	t
Business	2.44	.45	511	2.26*
Liberal Arts	2.53	.44		

* $p < .05$

Those students who declared business or education as a major did not predict GPA's significantl higher than those of students in any other majo field.

In order to discover the significance of the effe of various levels of the father's educational bac ground upon predictions of GPA, the data we subjected to the t-test.

Those students whose fathers' educational lev

Table 5.—t-Test Showing Significant Differences Between Predicted GPA's of Male Freshman Students Whose Fathers' Educational Level was Graduate Work and of Those Whose Fathers' Educational Level was 0–8 Years of School, High School, or College

GROUP N=1,025	MEAN	SD	NDF	t
0-8 Years of School	2.49	.43	311	4.17**
Graduate Work	2.72	.47		
High School	2.50	.48	491	3.85**
Graduate Work	2.72	.47		
College	2.53	.45	393	3.34**
Graduate Work	2.72	.47		

** $p < .01$

was graduate work predicted higher GPA's than did those whose father's educational level was 0–8 years of school, high school or high school graduate, or college or college graduate (see Table 5). Those students whose fathers' educational level was 0–8 years of school, high school or high school graduate, or college or college graduate did not predict GPA's significantly higher than those students whose fathers' educational level fell in any one of the listed categories.

Table 6.—t-Test Showing Significant Differences Between Predicted GPA's of Male Freshman Students Whose Average High School Grade was A and of Those Whose Average High School Grade was B, C, or D

GROUP N=1,025	MEAN	SD	NDF	t
Grade A	2.92	.52	503	4.71**
Grade B	2.61	.44		
Grade A	2.92	.52	542	8.23**
Grade C	2.41	.42		
Grade A	2.92	.52	82	3.68**
Grade D	2.46	.59		

** $p < .01$

The t-test applied to the data of average high school grades revealed that students whose average high school grade was A predicted higher GPA's than did those whose average grade was B, C, or D (see Table 6). Subjects whose average high school grade was B predicted higher GPA's than did those whose average was C, but not higher than did those whose average grade was D (see Table 7). Students whose average high school grade was C did not predict higher GPA's than did students whose grade was D.

Table 7.—t-Test Showing Significant Differences Between Predicted GPA's of Male Freshman Students Whose Average High School Grade was B and of Those Whose Average High School Grade was C

GROUP N=1,025	MEAN	SD	NDF	t
Grade B	2.61	.44	939	7.36
Grade C	2.41	.42		

** $p < .01$

Discussion

An investigation of selected demographic data as potential factors in male freshman students' predictions of GPA's revealed that ethnic background was not significant as a prediction determinant. There were significant effects on predicted GPA's when such factors as college major, socioeconomic level, and average high school grade were used as variables.

Students from upper socioeconomic levels did predict significantly higher GPA's than did students from lower socioeconomic levels; therefore, father's educational level had a direct effect upon student predictions, with the higher the father's educational level, the higher the GPA prediction.

Major area of college study and average high school grade also had a significant effect on GPA predictions. Science majors predicted significantly higher GPA's than did business majors, education majors, engineering majors, and liberal arts majors. However, those students who declared business or education as a major did not predict GPA's significantly higher than those of students in any other major field. Students who reported an average high school grade of A predicted higher GPA's than did those whose average grade was B, C, or D. Subjects whose average high school grade was B predicted higher GPA's than did those whose average was C, but not higher than those whose grade was D. However, students whose average high school grade was C did not predict higher GPA's than did students whose grade was B or D.

Of the factors considered, then, socioeconomic status, past experiences, and major area of study were significant factors in students' predictions of GPA.

Although a search of the related literature revealed no studies of both demographic factors and predicted GPA's, various studies of the effects of demographic factors on actual student performance have been conducted.

There is a definite implication for a follow-up study relative to this study in order to compare the same demographic effects on both predicted GPA and achieved GPA. To place this recommendation in the proper perspective, some pertinent studies which investigated the same demographic factors studied here are outlined briefly.

Bernstein (2) found that family size was inversely related to academic performance with the larger the number of siblings, the lower the level of school achievement. Nisbet (5) also found an inverse relationship between family size and intelligence.

Bradfield (3) investigated the effects of low-income backgrounds on male academic performance and found that, although students from such backgrounds showed personal characteristics similar to those which have appeared in college dropouts, they performed as well academically as other students. However, Wesley (8) stated that "high achievers appeared, when compared to low achievers, to have less concern about financial matters." Astin (1) found that the majority of dropouts from college came from the lower socioeconomic groups. Schroeder and Sledge (6) suggested that personal or motivational factors may be more important determinants of college achievement than socioeconomic level of parents.

Vaughan (7) conducted a study in which he compared Anglo-American and Latin-American students. The study concluded that Anglo-Americans score significantly higher than Latin-Americans on the American College Test and that scores on this test are significantly related to academic achievement. Records at the University of Texas at El Paso indicate that Mexican-American students generally do not rate as high academically as Anglo-American students. This information tends to indicate that students who are from lower socioeconomic backgrounds and have learned English as a second language do not perform as well as other students.

Implications

There are particular implications for student financial aid departments and for academic advising. Since economic factors were significant in student prediction of academic success, with students from lower economic backgrounds predic ing lower, it may be hypothesized that predi tion would be affected by freedom from econom worry. Many students find it necessary to wo while attending college, thus depriving themselv of time for study and leisure and lowering the chances for academic success. Financial aid pr vided for such students would not only relie economic problems, but would also permit add tional study time and enhance their opportuniti for scholastic achievement in college.

More intelligent direction would be provide through the counselor's use of trends revealed i this investigation and his knowledge of the student's background and level of aspiration. Keefe (4) found that self-prediction of academic success was a stable and significant variable which proved to be as reliable as the high school recor and standardized achievement test results in the prediction of academic success. In some instances, the self-prediction was found to be a more accurate predictor. There was a tendency for the high school record and standardized achievement tes score to lose predictive accuracy as a studen moved away from the freshman year, while the self-prediction retained its accuracy at essentially the same level.

To investigate in depth the results of the stud with male freshman students, a follow-up stud is now in progress to obtain the achieved freshman GPA's for comparison with the predicte GPA's in relation to the demographic variable under consideration. Additional areas of signif cance may be revealed when the factor of achieve GPA is introduced.

On the basis of the results of this study, furthe action research is recommended. The following i a description of such a study. Two groups woul be formed on a matching basis. An experiment group of freshman candidates would be identifie on the basis of low socioeconomic level, low ave age high school grade, and college majors chose by students with typically low aspirational leve as measured by the current study; a control grou with the same characteristics would then formed. Since ethnic background revealed no si nificant effect on prediction of GPA, this fact would not be considered. Both the experiment group and the control group would be asked predict their GPA's for the first semester of the freshman year in college.

The experimental group would be brought t college campus during the summer immediat prior to entrance as freshman students. These s dents would receive an intensive summer cou in study skills, and they would experience the ing conditions of the college environment. Th would be exposed to professors, guest lectur and other students attending the summer sess

At the end of the summer both groups would again be asked to predict their GPA for the first semester. The results would provide information as to whether the change in prior conditions of the students in the experimental group had a significant effect on their aspiration levels. If a significant effect was found, a comparison of achieved GPA's of the two groups at the end of the first semester would be made to validate the findings. It could then be recommended that these "special" students be identified and enrolled in an orientation session for the purpose of improving their general outlook and subsequent performance.

REFERENCES

1. Astin, A. W., "Personal and Environmental Factors Associated with College Dropouts among High Aptitude Students," *Journal of Educational Psychology*, 55: 219–227, 1964.
2. Bernstein, B., "Some Sociological Determinants of Perception: An Inquiry into Sub-cultural Differences," *British Journal of Sociology*, 9:159–174, 1958.
3. Bradfield, L. E., "College Adjustment and Performance of Low-income Freshman Males," *Personnel and Guidance Journal*, 46:123–129, October 1967.
4. Keefer, K. E., "Self-prediction of Academic Achievement of College Students," doctorial dissertation, University of Tennessee, University Microfilms, Ann Arbor, Michigan, 1966, No. 66–178.
5. Nisbet, J., "Family Environment and Intelligence," in Halsey, A. H.; Floud, J.; Anderson, C. A. (eds.), *Education, Economy, and Society*, The Free Press, New York, 1961, pp. 273–287.
6. Schroeder, W. L.; Sledge, G. W., "Factors Related to Collegiate Academic Success," *Journal of College Student Personnel*, 7:97–104, 1966.
7. Vaughan, G. E., Jr., "Some Characteristics of College Freshmen According to Sex and Ethnic Group and the Relationship of these Characteristics to Academic Achievement," doctoral dissertation, North Dakota State University, University Microfilms, Ann Arbor, Michigan, 1964, No. 64–5183.
8. Wesley, D., "The Relationship between Psychosocial Factors and Academic Achievement of Selected College Freshmen," doctoral dissertation, Oklahoma State University, University Microfilms, Ann Arbor, Michigan, 1962, No. 62–1628.

New Books from the Publishers

Children

Child Study Association of America. *Children's Books of the Year 1969* (New York: Child Study Association of America, 1970), 48 pp. Paper $1.50.

Evans, Judith L. *Children in Africa: A Review of Psychological Research* (New York: Teachers College Press, 1970), 115 pp. Paper $3.95.

Lessinger, Leon. *Every Kid a Winner—Accountability in Education* (New York: Simon and Schuster, Inc., 1970), xv + 239 pp. Hardcover $7.95.

O'Gorman, Ned. *The Storefront: A Community of Children on 129th Street and Madison Avenue* (New York: Harper and Row, Publishers, 1971). Hardcover $4.50.

Wright, Betty Atwell and others. *Elementary School Curriculum: Better Teaching Now* (New York: The Macmillan Company, 1971), xi + 403 pp. Hardcover $8.95.

College

Daiches, David (ed.). *The Idea of a New University* (Cambridge, MA: The MIT Press, 1971). Paper $2.95. Cloth $10.00.

Davis, Martin W. *The Sister as Campus Minister* (Washington, D.C.: Center for Applied Research in the Apostolate-CARA, 1970), xvi + 124 pp. Paper.

Horn, Francis H. *Challenge and Perspective in Higher Education* (Illinois: Southern Illinois University Press, 1971), xvi + 224 pp. Hardcover $8.95.

McLeish, John. *Students' Attitudes and College Environments* (Cambridge, England: Cambridge Institute of Education, 1970), vii + 251 pp. Paper $6.00.

Pitcher, Robert W. and Babette Blaushild. *Why College Students Fail* (New York: Funk and Wagnalls, 1970), vi + 271 pp. Hardcover $6.95.

Sedgwick, Adam. *A Discourse on the Studies of the University* (New York: Humanities Press, 1970), 109 pp. Hardcover $5.00.

Simms, Theodore F. *Improving College Study Skills: A Guide and Workbook* (Beverly Hills, CA: Glencoe Press, 1970), viii + 159 pp. Paper $3.75.

Current Issues

Burridge Trevor. D. *What Happened in Education: An Introduction to Western Educational History* (Boston, MA: Allyn and Bacon, 1970), v + 146 pp. Paper $3.95.

Eide, Ingrid. *Students as Links Between Cultures* (Boston, MA: Universitetsforlaget, 1970), ix + 243 pp. Hardcover $10.00.

Ellsworth, Frank L. and Martha A. Burns. *STUDENT ACTIVISM in America Higher Education, Student Personnel Series No. 10* (Washington, D.C.: American College Personnel Association, Division of American Personnel and Guidance Association, 1970), 64 pp. Paper $2.50.

Sudbury Valley School (The). *The Crisis in American Education* (Framingham, MA: The Sudbury Valley School Press, 1970), 109 pp. Paper $1.50.

Education

Armstrong, R. H. R. and John L. Taylor (eds.). *Instructional Simulation Systems in Higher Education* (Cambridge, England: Cambridge Institute of Education, 1970), v + 216 pp. Paper.

Battle, Vincent M. and Charles H. Lyons (eds.). *Essays in the History of African Education* (New York: Teachers College Press, 1970), ix + 123 pp. Paper $3.50.

Bourne, Randolph S. *The Gary Schools* (Cambridge, MA: The MIT Press, 1970), xvii + 323 pp. Hardcover $12.50.

Cleugh, M. F. *Educating Older People* (New York: Barnes and Noble, Inc., 1970), 187 pp. Paper $3.00.

Eisner, Elliot W. *Confronting Curriculum Reform* (Boston, MA: Little, Brown and Company, 1970), xi + 206 pp. Hardcover $7.00.

Gustafson, James M. and others. *Moral Education: Five Lectures* (Cambridge, MA: Harvard University Press, 1970), 136 pp. Hardcover $4.95.

Hillesheim, James W. (ed.). *Theory and Practice in the History of American Education* (Palisades, CA: Prentice-Hall, 1971), xx + 422 pp. Cloth $9.95.

van Kreveld, D. *Structure and Outcome of Study Groups* (Netherlands: Koninklinjke van Gorcum and Company, Inc., 1970), 75 pp. Paper Hfl. 9, 50 Sewn.

Lawlor, Elizabeth Phelan. *Research in Science Education 1953 through 1957* (New York: Teachers College Press, 1970), viii + 112 pp. Hardcover.

McNeill, Malvina Rosat. *Guidelines to Problems of Education in Brazil* (New York: Teachers College Press, 1970), vi + 66 pp. Cloth $3.95.

Samuels, Michael Anthony. *Education in Angola, 1878–1914: A History of Culture Transfer and Administration* (New York: Teachers College Press, 1970), xiii + 185 pp. Paperback $7.95.

NAL OF EDUCATIONAL RESEARCH
4, Number 9, May–June 1971)

Sex Differences in the School Gains of Migrant Children

SUSHILA SINGHAL and PRISCILLA H. CRAGO
New York State Education Department

ABSTRACT

The purpose of this study was to examine the effects of sex differences in the school gains of migrant children. Results from the Wide Range Achievement Test showed significant differences in the reading achievement of boys and girls at the grade 4 level in favor of boys. In arithmetic achievement, differences were significant at the grade levels 3, 4, and 9 in favor of boys. In the total group, boys and girls did not differ in reading or arithmetic gains. No indications were found of migrant children's school gains being affected adversely either by the teachers' discriminatory behavior in favor of girls or by the family structure. However, the present research does indicate the need for replicating the study to explain the differential growth in learning noted at certain grade levels within migrant and other disadvantaged groups.

THAT GIRLS demonstrate higher achievement in reading and boys demonstrate higher achievement in arithmetic has been the belief of most classroom teachers and educators. Lincoln (4) found that in reading at the elementary school level girls tended to excel by small margins. Stroud and Lindquist (8), using the Iowa Every Pupil Basic Skills Testing Program, reported a consistent and significant superiority in language, work study, reading, vocabulary, and verbal comprehension for girls, while small insignificant differences in arithmetic favored boys. Gates (2) confirmed that on the average girls' reading abilities excel boys' in grades 2 through 8. Durrell (1) and Newton and others (5) reported that more boys attend reading clinics than do girls. However, Kowitz and Mahoney (3), Parsley and others (6), and Sinks and Powell (7) have reported insignificant sex differences in their studies.

While these studies have examined sex differences in achievement status at several grade levels, very few studies have investigated sex differences in growth rate. The present study was designed to examine the following hypothesis empirically:

> *Boys and girls from the migrant families do not differ in their school gains in reading and arithmetic.*

Method

Subjects. Included in the study were 777 boys and 775 girls from migrant workers' families in New York State for whom pretest and posttest data were complete. These children do not have opportunities for regular schooling, and are culturally and economically disadvantaged. In the summer of 1969, thirty-five migrant centers were functioning in New York State to insure continued pupil growth by providing compensatory educational opportunities to these children. Their ages ranged between 5 and 16 years.

Test. The Wide Range Achievement Test (WRAT) (1965 Edition, Level I) was selected as a measure of migrant children's achievement in reading and arithmetic at the beginning and end of the program. The test provides for a wide range scaling and is applicable to a heterogeneous age group (from kindergarten to college). The test has a high reliability, of .93, with subtest reliability coefficients ranging between .90 and .95. The test yields grade-equivalent scores, which provide the teachers with bases for choosing appropriate instructional materials for each child at the start of the program.

Procedure. Each child in the group was tested at the beginning and end of the program, which lasted from early July to the middle or end of August, 1969. The obtained raw scores were converted to grade-equivalents using the test publisher's table. The group was classified by age ranges corresponding to the usual grade placement. For example, children between 4 years, 8 months and 5 years, 7 months in July would be between 4 years, 10 months and 5 years, 9 months in September, approximately the expected school age range for admission to kindergarten.

Statistical analysis was done for the total group and for the breakdown by expected grade level. Mean gains between pretest and posttest data were computed for boys and girls separately. Student t's for independent samples were computed.

Results and Discussion

Girls obtained higher mean grade-equivalent scores on the WRAT reading and arithmetic pretests than boys, as a total group and at most grade levels. This finding partially supports past studies which report higher reading achievement for girls.

Considering the total gains of each sex group by itself, girls gained more in reading (.27 grade-equivalents or 2.7 months) than they did in arithmetic (.19), and boys gained more in arithmetic (.34) than they did in reading (.30). By grades, however, boys gained more in arithmetic than in reading only in grades 2 and 3, while girls gained more in reading than in arithmetic in grades 4 through 10. Thus, the mean gains in reading were higher than the mean gains in arithmetic for both sexes in grades 4 through 10. The achievement in arithmetic required by WRAT above the third-grade level involves the development of abilities to solve oral problems and perform written computations—skills which develop gradually, while the skills called for from kindergarten through grade 3 involve the relatively easy development of the abilities to count and read number symbols. The nature of the tasks may have made both sexes gain more in arithmetic than in reading in the lower grades. Also, the slow arithmetic gains above the third-grade level might have occurred as a result of the regression effect. This tends to support the hypothesis of no difference in gains in reading and arithmetic between boys and girls. Because the migrant children were in the program for only 6 to 8 weeks, they did not acquire the computational skills to the required level, although they developed abilities in recognizing letters, naming letters, and pronouncing words. Nevertheless, the expected gain in arithmetic for a 6- to 8-week period of .15 to .20 grade-equivalents was met by the girls who showed a .19 gain and exceeded by the boys who made a .34 gain.

When the significance of the sex differential of gains is tested (see Table 1), it is found that reading gains were significant in favor of boys in grade 4 ($p < .05$) and tended toward significance in favor of boys in kindergarten ($p < .10$). The arithmetic gains, as well, were significantly higher for boys than girls in grades 3, 4, and 9 ($p < .05$). In no case did the superior gains for girls (in grades 5 and 6 reading and in kindergarten, grades 7, 10, and 11 arithmetic) reach

Table 1.—Mean Gains, Gain Differences, and t-values in WRAT Reading and Arithmetic Grade-Equivalents for Boys and Girls

Grades	Boys			Girls			Reading		Arithmetic	
	N	Mean Gains Rdg.	Mean Gains Arith.	N	Mean Gains Rdg.	Mean Gains Arith.	Gain Differ-ences	t-value	Gain Differ-ences	t-value
KG	53	.48	.27	54	.25	.33	.23	1.53 b	.06	.46 g
1	90	.27	.25	105	.15	.19	.12	1.09 b	.06	.86 b
2	120	.12	.40	129	.11	.21	.01	.13 b	.11	.79 b
3	129	.12	.49	85	.19	.22	.07	.58 g	.27	1.93*b
4	123	.35	.26	105	.14	.05	.21	2.10 *b	.21	2.63**b
5	81	.25	.21	89	.53	.12	.28	1.22 g	.09	.60 b
6	76	.33	.22	84	.44	.20	.11	.35 g	.02	.22 b
7	45	.54	.33	52	.41	.37	.13	.50 b	.04	.24 g
8	32	.50	.26	46	.37	.07	.13	.48 b	.19	.59 b
9	17	.88	.61	14	.63	.10	.25	.38 b	.51	2.32*b
10	6	2.29	.52	8	1.20	.88	1.09	.78 b	.36	.38 g
11	5	.28	.00	4	.15	1.18	.13	.09 b	1.18	.86 g
Total	777	.30	.34	775	.27	.19	.03	.07 b	.15	.47 b

*$p < .05$ in favor of boys
**$p < .01$ in favor of boys
b—nonsignificant difference favoring boys
g—nonsignificant difference favoring girls

significance at the .05 level. Nor were there any significant differences in the gains for one sex over the other for the total groups. The hypothesis of no differential gains of statistical significance is accepted, therefore, for the total groups but not for selected grade levels. The fact that any significant gains noted favored the boys, however, is of interest. The gains occurred at three grade-levels in arithmetic, as might be expected and at only one grade-level in reading.

Results did not indicate that girls were likely to identify with the teachers and consequently do better in school. Nor were there indications of teachers' behavior being discriminatory toward boys or girls. Although many of the migrant families happened to be mother-dominated, the school achievement of boys was not adversely affected.

The migrant centers did contribute to the cognitive development of migrant children effectively. The instructional setting provided by the teachers and other school staff was very relevant to the needs of this group without considering the sex. The present findings, though meager and inadequate, do indicate the need of replication of this research. Attempts should be made to find more direct explanations of differences in rates of growth and achievement within the context of programs for disadvantaged children such as those from migrant families. The teachers may be instructed to record their occasional observations of each child's reactions to the classroom situations and materials included in the program.

REFERENCES

1. Durrell, D. D., "Improving Reading Instruction," World Book Company, Chicago, Illinois, 1956.
2. Gates, Arthur I., "Sex Differences in Reading Ability," *The Elementary School Journal*, 61:431–434, 1961.
3. Kowitz, Gerald T.; Mahoney, Leo G., "Sex Linked Factors in Reading and Arithmetic Achievement," paper presented at the American Educational Research Association Convention, Chicago, Illinois, February 10–14, 1964 (mimeo).
4. Lincoln, E. A., *Sex Differences in the Growth of American School Children*, Warwick and York, Baltimore, Maryland, 1927.
5. Newton, Eunice Shaed and others, "Empirical Differences Between Adequate and Retarded Readers," *The Reading Teacher*, 13:40–44, 1959.
6. Parsley, Kenneth M.; Powell, Marvin; O'Connor, Henry; Deutch, Murray, "Are There Really Sex Differences in Achievement?" *The Journal of Educational Research*, 57:210–212, 1963.
7. Sinks, Naomi B.; Powell, Marvin, "Sex and Intelligence as Factors in Achievement in Reading in Grades 4 Through 8," *Journal of Genetic Psychology*, 106:67–79, 1965.
8. Stroud, J. B.; Lindquist, E. F., "Sex Differences in Achievement in the Elementary and Secondary Schools," *Journal of Educational Psychology*, 33:657–667, 1942.

Forthcoming Articles

Borup, Jerry H.; Weber State College, Ogden, Utah
"THE VALIDITY OF AMERICAN COLLEGE TEST FOR DISCERNING POTENTIAL ACADEMIC ACHIEVEMENT LEVELS—ETHNIC AND SEX GROUPS"

DeBlassie, Richard R.; New Mexico State University, Las Cruces, and W. Paul Jones; Science Research Associates, Chicago, Illinois
"SOCIAL CLASS CONTRASTS IN SHORT TERM PREDICTABILITY OF GRADE 7 ACHIEVEMENT"

Gamsky, Neal R. and Faye Williams Lloyd; Cooperative Educational Service, Waupun, Wisconsin
"A LONGITUDINAL STUDY OF VISUAL PERCEPTUAL TRAINING AND READING ACHIEVEMENT"

Gies, Frederick John, B. Charles Leonard, and Jon S. Paden; Center for Educational Improvement, University of Missouri—Columbia
"THE EFFECT OF SELECTED MEDIA FEEDBACK UPON THE INTERACTIVE BEHAVIOR OF STUDENT TEACHERS"

Grant, Claude W. and Lois H. Worthington; University of Utah, Salt Lake City
"FACTORS OF ACADEMIC SUCCESS: A MULTIVARIATE ANALYSIS"

Haddad, Wadi'D.; American University of Beirut, Lebanon
"NATURE AND LANGUAGE OF SCIENTIFIC TERMS IN THE ARABIC PRESS OF THE MIDDLE EAST"

Stein, Sandra Lou; University of South Carolina, Columbia
"THE INTERRELATIONSHIPS AMONG SELF-ESTEEM, PERSONAL VALUES, AND INTERPERSONAL VALUES"

Thorman, Joseph H.; Moorhead State College, Minnesota
"RELATIVE EFFECTIVENESS OF FOUR METHODS OF TRAINING PROSPECTIVE TEACHERS IN INTERPERSONAL SKILLS"

Walters, Donald L.; Temple University, Philadelphia and Ralph L. Spencer, State University of New York at Oswego
"A STUDY OF ATTITUDES TOWARD NEGATIVE ADMINISTRATIVE STEREOTYPES"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

The Effect of Grade Contracts on Student Performance[1]

WILLIAM A. POPPEN and CHARLES L. THOMPSON
The University of Tennessee

ABSTRACT

Each of four classes of Educational Psychology 2430(Child Study) was randomly divided into two groups: experimental (N=55) and control (N=55). The experimental group's course performance was evaluated by a grade contract approach while the control group was evaluated by a traditional method. Group differences favored the experimental group; however, such differences were not statistically significant.

RECENTLY MUCH criticism has been leveled at traditional approaches to the evaluation and grading process utilized in higher education. While campus unrest may attest to student dissatisfaction with many university policies including grading, the problem has been clearly identified in the literature for a number of years. Over 50 years ago, Rugg (4) pointed out the lack of uniformity and the questionable validity of a grading system by which is determined such crucial matters as promotion, college admissions, graduation, and selection for honoraries, scholarships, and fellowships. Miller's (2) recent survey of studies on grading tend to bear out the fact that grading practices still seem to suffer from low validity and reliability. From his review, he concluded that present grading practices tend to discourage those qualities which we generally attribute to scholarly inquiry: Creativity and independent thought. Hoyt's (1) review of forty-six grading studies illustrated the problem of low validity in grading practices. Most of the studies reviewed found little relationship between college grades and various success measures in the "outside" world.

The obvious conflicts existing between the questionable validity of traditional grading practices and their influences upon learning and career development have not stimulated a great deal of innovation in our systems of learning evaluation. Most innovative attempts have been variations of the two-letter pass–fail system (2). Frequently these variations of the traditional A, B, C, D, and F system amount to little more than symbol manipulation.

Wrinkle (6), reporting on the most popular grading deviation (the pass–fail or two-symbol system), found that many students were only concerned with staying above the fail limit. A further disadvantage was found in the lack of information conveyed by the pass–fail marking system.

Clearly, a need exists for a system of student evaluation which does not impede learning or motivation and still provides accurate feedback to the student. The study reported herein represents one attempt to fill the credibility gap in grading.

Method

Each of four classes of Educational Psychology 2430 was randomly divided into two groups: experimental (N=55) and control (N=55). The experimental group's course performance was evaluated by the grade contract approach while the control group's performance was evaluated by a traditional method.

The Grade Contract Method. The grade contract (5) is designed to give the student choices in selecting his learning activities as well as the grade level he wishes to achieve. The contract provides a structural continuum from complete structure to complete freedom for the student in selecting his learning activities. He may choose to meet suggested grade level requirements or he may substitute individual or independent learning activities for any or all of the suggested course requirements. Substitution value of learning activities is a manner of negotiation between instructor and student at the contract signing time. The contract is renegotiable at any time during the quarter if the student feels he is not able to meet contract terms. Failure to meet contract terms simply means that the student qualifies for the grade representing the level achieved whether this be above or below the contracted level. A student contracting for a "C" grade and achieving a "B" grade would naturally be awarded a "B."

The grade contract has many advantages for both student and instructor. Student anxiety about grades is diminished as ambiguity about course requirements is canceled by the contract's explicitness. Motivation to learn should be increased as the student becomes more involved in the self-direction of his learning and self-evaluation of his progress. Suddenly he finds himself in an academic course where he has choices and an opportunity to behave responsibly. The student has an opportunity to succeed at his chosen level.

The instructor also benefits from the grade contract. Course objectives must be clarified before the contract can be implemented. Clarification of learning objectives also removes much of the ambiguity in grading student performance. Grading becomes a matter of determining a pass-fail grade for the student's contracted grade level rather than assigning grades of A, B, C, D, and F. Course grading curves go out the window because each student is doing his thing. An overall advantage seems to be the instructor–student learning partnership which can emerge from the contract method (see Appendix A for a copy of the grade contract used in the study).

The Traditional Grading Method. For purposes of the study, a traditional method of grading was used for the control groups. It involved computing an average grade from scores made on two standardized tests, a case study report, and an outside reading report. Students were given no choices in their assignments and all students were evaluated on the same work, even following the same general outline on the case study report of the child study phase of the course. The instructors did mention that independent projects could be done by the students for extra grade credit.

Evaluation of Criteria. The effect of the grade contract upon student learning was assessed by the following procedures:

1. *Comparison of midterm and final examination scores for each group (both examinations are standardized tests consisting of true-false and multiple-choice items).*
2. *Comparison by raters (three doctoral students in Educational Psychology were employed as raters) of the quality of child case study reports completed by the members of each group.*
3. *Comparison by raters of the quality and number of reaction papers completed by the members of each group.*
4. *Comparison by raters of the quality and number of independent projects completed by the members of each group.*
5. *Comparison of student responses to a semantic differential form (the semantic differential (see Appendix B) is a technique developed by Osgood and others (3) for measuring concept meaning by having Ss respond to a bipolar adjective checklist) containing the following concepts: (a) Educational Psychology 2430; (b) Case Study; (c) Outside Reading; (d) Grading Procedure; (e) Independent Study.*

Treatment of Data. Data collected from the test results, raters, and the semantic differential were evaluated for group differences by an analysis of variance.

Findings

Table 1 presents a comparison of grades and test scores obtained by each of the treatment groups. While each of the four means favor the treatment (grade contract) group, none of the F values are significant at the .05 level of significance.

Table 1.—Mean Grades and F ratios for the Treatment and Control Groups on Course Assignments

Variable	Grade Contract Mean	Non-Contract Mean	F
Final Grade*	2.964	2.911	0.182
Case Study Grade*	3.009	3.045	0.099
Character Analysis*	2.946	2.938	0.004
Two Test Scores**	148.875	146.946	0.448

F of 3.92 needed for .05 level of significance

*Based on a 4-point scale: A=4, B=3, etc.

**Total raw scores for the two examinations: Possible total score=200.

Table 2 presents a comparison of the means for various concepts of the semantic differential evaluative factor. The evaluative factor reportedly assesses the "goodness" or "badness" dimension of concept. In this particular situation the highest possible evaluation score was 21 and the lowest score 7. Accordingly, the scores which ranged from 15 to 18, are considered indicative of positive attitude rather than negative ratings

The data presented in Table 3 show the means for both treatment groups on the semantic differential total score. The mean scores for the grade contract group are higher for all four concepts; however, none are significant at the .05 level of significance.

Summary

The lack of significant difference between the

Table 2.—Mean Concept Ratings and F ratios for the Semantic Differential Evaluative Factor for the Treatment and Control Groups

Variable Evaluative Factor	Grade Contract Mean	Non-Contract Mean	F
Semantic Differential Concept			
Course	18.268	17.250	2.365
Case Study	18.929	18.054	2.021
Grading Procedures	15.821	15.536	0.119
Outside Readings	16.464	15.518	1.057

F of 3.92 needed for .05 level of significance

Table 3.—Mean Concept Ratings and F ratios for the Semantic Differential (Total Factor) for the Treatment and Control Groups

Variable Total Factor	Grade Contract Mean	Non-Contract Mean	F
Semantic Differential Concept			
Course	47.089	44.875	2.622
Case Study	49.750	47.018	3.460
Grading Procedures	44.929	44.036	0.191
Outside Readings	45.464	43.357	1.439

F of 3.92 needed for .05 level of significance

treatment and control groups leads the experimenter to draw one of two possible conclusions: (1) our traditional approach is not all that bad, or (2) our grade contract approach is not all that good.

Sallying forth beyond the data into a world of introspection produced the following thoughts:

1. *Perhaps the grade contract works best for students who are in academic trouble. Previous studies (5), have indicated considerable success with grade contracts for students experiencing academic failure. Average grades for Educational Psychology generally fall around the C+ mark.*
2. *Some independent projects stimulated by the grade contract approach were valued highly as learning experiences by the students and instructors. At least one research paper will be submitted for publication.*
3. *The two multiple-choice "objective" examinations are not contributing much to student learning. In fact, the tests may be hindering independent study. Subjective evaluation by the instructors indicates that those students opting independent study projects in place of the examinations became more involved in "relevant" learning.*
4. *Revision of the grade contract is needed to provide further for individual student differences and needs. For example, an option is needed which allows the student to choose an "incomplete" in the course if he is not able to meet his contract terms by the end of the quarter. The student should be able to choose between an "incomplete" or the grade for which he has qualified if he has not met his contract terms.*

Finally, comments by some of the students give an indication of the grade contract's possible worth.

Grade Contracts:

1. "Help students to know what is expected of them to meet course requirements."
2. ". . . provide for individual differences."
3. ". . . shift responsibility of 'becoming educated' to the student."
4. ". . . help the instructor to clarify course objectives."
5. ". . . give the students a voice and a choice."
6. ". . . build cooperation between student and instructor."

Conclusions

Are the grade contracts of value? This research does not substantiate the value of grade contracts over traditional approaches when used with college students. There may be undetected value or value that may be obtained with further modifications. The procedure certainly does allow movement toward individualization of instruction and seems to have no real disadvantages as a grading procedure.

FOOTNOTE

1. The research was supported by a University of Tennessee Mini-Grant.

REFERENCES

1. Hoyt, D., "The Relationship Between College Grades and Adult Achievement: A Review of the Literature," *Research Reports*, American College Testing Program, 7:i, 1965.
2. Miller, S., *Measure, Number, and Weight: A Polemical Statement of the College Grading Problem*, Learning Resources Center, The University of Tennessee, Knoxville, 1967.
3. Osgood, C.; Suci, G.; Tannenbaum, P., *The Measurement of Meaning*, University of Illinois Press, 1957.
4. Rugg, H., "Teacher's Marks and Marking Systems," *Educational Administration and Supervision*, 1:117–142, 1915.
5. Thompson, C.; Davis, M., "Grade Contracts: A Method of Redirecting Motivation," *Focus on Guidance*, in press.
6. Wrinkle, W., *Improving Marking and Reporting Practices*, Rinehart and Company, Inc., New York, 1963.

APPENDIX A

Educational Psychology 2430
Grade Contract

Grade A

1. Score at least 85% on the midterm and final exams.
2. Complete an "A" level case study report (see case study guidelines).
3. Do a before and after sociometric analysis of your case study child's class and report the results in the case report.
4. Do the student problem survey with your case study child's teacher.
5. Attend classes and quiz sections a minimum of twenty-three times *(roll will be taken)*.
6. Participate in classes and quiz sections.
7. Conduct a home visit and an outside observation experience with your child; report these data in your case write-up.
8. Turn in your daily observation notes at the end of the quarter.
9. Observe your child at least seven times during the quarter.

Grade B

1. Score at least 75% on the midterm and final exams.
2. Complete a "B" level case study report (see case study guidelines).
3. Do a before and after sociometric analysis of your case study child's class and report the results in the case report.
4. Do the student problem survey with your case study child's teacher.
5. Attend class and quiz sections a minimum of twenty-three times *(roll will be taken)*.
6. Participate in classes and quiz sections.
7. Conduct a home visit and an outside observation experience with your child; report these data in your case study write-up.
8. Turn in your daily observation notes at the end of the quarter.
9. Observe your child at least seven times during the quarter.

Grade C

1. Score at least 65% on the midterm and final exams.
2. Complete a "C" level case study report (see case study guidelines).
3. Do a before and after sociometric analysis of your case study child's class and report the results in the case report.
4. Do the student problem survey with your case study child's teacher.
5. Attend classes and quiz sections a minimum of twenty-three times *(roll will be taken)*.
6. Participate in classes and quiz sections.
7. Conduct either a home visit *or* an outside observation experience with your child.
8. Turn in your daily observation notes at the end of the quarter.
9. Observe case study child at least six times during the quarter.

Grade D

1. Score at least 55% on the midterm and final exams.
2. Observe case study child at least six times during the quarter.
3. Turn in your daily observation notes at the end of the quarter.
4. Attend classes and quiz sections a minimum of twenty-three times *(roll will be taken)*.
5. Conduct *either* a home visit *or* an outside observation experience with your child.

Grade F

1. Nothing for nothing.

---

* An Independent Project of your design may be substituted with instructor approval for *any* or *all* of the above requirements. Use the Independent Project form to indicate the nature of your project.

Grade Contract

I have read the attached description of the requirements for the grades A, B, C, D, and F and on the basis of these requirements I, ________________________________, wish to contract for grade of _________. If I fail to fulfill the requirements for my chosen grade, I will accept the grade for which I have qualified. I understand that I have the option to renegotiate this contract during the week following midterm examinations.

Date ____________________ Student ________________

Renegotiation ______________ Instructor ______________

Individualized Contract Terms

1.
2.
3.
4.

INDEPENDENT STUDY

Course number ______ Term ______ Student ____________

Independent study can be done in lieu of any class requirement, such as a term paper, assigned project, or examination. The scope and depth of the study depends upon the requirements it is replacing. The process of proposing independent study can be accomplished by filling out the following outline.

1. What do you want to find out?
2. How are you going to do it (resources and procedures)?
3. How do you plan to evaluate and report your learning?
4. What grade requirement(s) does this study replace?

USE THE BACK OF THIS SHEET IF ADDITIONAL SPACE IS NEEDED

I certify that this study is a new learning experience for me. ------------------------------------

Deadline for submission of proposal ----------------

Deadline for resubmission -----------------------

Instructor's approval ---------------------------

APPENDIX B

Semantic Differential Form

*Concept

Strong :___:___:___:___:___:___:___: Weak

Deep :___:___:___:___:___:___:___: Shallow

Bad :___:___:___:___:___:___:___: Good

Narrow :___:___:___:___:___:___:___: Wide

Positive :___:___:___:___:___:___:___: Negative

Complicated :___:___:___:___:___:___:___: Simple

Predictable :___:___:___:___:___:___:___: Unpredictable

Understandable :___:___:___:___:___:___:___: Mysterious

Worthless :___:___:___:___:___:___:___: Valuable

---

* The concepts used in this study were: Educational Psychology 2430, Case Study, Outside Reading, Grading Procedure, and Independent Study.

Book Reviews

ROBERT E. CLASEN
Book Review Editor

Issues in American Education

Kroll, Arthur M. and others. Oxford University Press, England, 1970, 202 pp. $6.00.

ONE WAY to review a book is to report some of the ideas and questions the book engenders in the mind of the reviewer. Since this method has been adopted, the authors should not be held responsible for the mental meanderings of the reviewer. The following report of the reviewer's mental williewaws is a way of saying *Issues in American Education* is a very stimulating book. Several times while reading the book's treatment of national assessment, technology in education, vocational guidance, desegregation, and other topics, the reviewer said to himself—*brilliant and innovative.*

The book discusses values. Yes, values change, but toward what do they change? What is the anchor value? How does one test values? What is the correct philosophical ultimate? The Creator endowed each of us with guides, which are the most authentic revelations. Moses should have brought down an eleventh commandment: *Thou shalt satisfy the urges with which I have endowed thee.*

It is strange how men rear back from acknowledging the truth visible in their every action. They even contend that man should not aim at satisfaction—but that instead he should hope that it will be a by-product of some presumed nobler aim. These aims thereby become fixed, they create generation gaps and block progress toward man's true goal in education.

A young lady claimed that the ultimate good was the greatest good of the greatest number until she was reminded that she had just married the young man a dozen other girls wanted. Discussion of goals of education tends to be sterile until it gets down to consideration of the real dynamics of life, namely, an individual's urges and his efforts to satisfy the urges which inhabit him.

Here are some issues in education which should be posed. Since out of the heart are the issues of life and a student's urges are the most intimate portion of his personality, to what extent is it permissible for a school to alter urges? Are educators in error when they teach a student to be unselfish and sacrifice his urges to the urges of others especially since others wear so many masks that a student's estimate of what will satisfy them is much less accurate than estimates he makes for himself? How far ahead should a student make his estimates? Does the amount of satisfaction alter with time?

Goals of education should not be the direct teaching of either facts or principles. Textbooks having such content, as they often do in science and social studies, should be placed on the reference shelf. Only skills such as reading, reasoning, composition, and mathematics should be the direct aims of instruction, and development of these skills should be by-products.

How about values? Sensible adults learn facts and principles while trying to satisfy some personal urge. Students from the first grade through the university should be no less sensible.

The book under review discusses the national assessment test. Despite the reviewer's general approval of the national assessment test, he was amazed to find the test containing such an insignificant factual item as the metal used in making tin cans. It would have been entirely appropriate to test whether a student knew how to go about finding the answer to such an item. It is a tribute to Dr. Tyler's genius that he did so well in view of all the groups who had to be satisfied.

The book lists art, music, and literature as proper goals of education. In every school there should be a large amount of participation, sharing, and enjoyment of all three. But it should all be in fun. The supervisor of music kindly guided me to a music class. When I asked the children to sing for me, the teacher directed them to turn to a certain page in the music book. The pupils-struggled-through-the-song. Then this reviewer asked them to sing a song of their own choice. They chose "Don't Fence Me In" and nearly raised the roof. Except for incidental teaching while students are enjoying music or other arts and except for the few with exceptional interest, there should be no set courses with achievement marks. This reviewer has heard more singing coming from Smoky Mountain coves than he has heard coming from inhabitants in urban or suburban homes.

William A. McCall
Emeritus Professor of Education
Teachers College, Columbia University

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 9, May–June 1971)

Sociometric Predictors of the Self Concept Among Fifth and Sixth Grade Children[1]

BERT O. RICHMOND and WILLIAM F. WHITE
The University of Georgia

ABSTRACT

Theories of self-concept usually maintain that the individual's self-esteem is reflected in peer ratings. The purpose of this study is to isolate factors of self-esteem and of peer ratings and to determine significant relationships between the derived factors. Over two hundred elementary pupils were selected from two metropolitan areas. Approximately half were black, disadvantaged pupils; the remainder were white with above average cultural advantages. The Coopersmith Self-Esteem Inventory (SEI) and a semantic differential (SD) were administered to all students. From a factor analysis of responses to the SEI, four factors of self-esteem and a lie scale emerged. Three SD factors were obtained from a factor analysis of the peer ratings. One significant canonical correlation (.325 $p < .01$) resulted from the correlation of the three SD factor scores with scores on the five SEI factors. "Activity" dominated the relationship between peer feelings and self-concept dimensions.

MUCH RECENT educational research places emphasis on understanding the relationship between scholastic ability and self-concept and other psychological characteristics of the individual. Studies by Brookover (1), Combs (2), Gordon (4), Rentz and White (5), and Wylie (6) represent some of the attempts to explore these relationships in depth. Coopersmith's (3) book-length report of research with the Self-Esteem Inventory (SEI) is a comprehensive effort to define antecedents of self-esteem. He uses the SEI, which has five theoretically derived factors of self concept, together with selected rating scales to provide information about the social behavior of children with low, middle, and high self-esteem. One of his conclusions is that persons of low self-esteem are likely to be less visible members of the crowd than group leaders and thus more limited in social intercourse with decreased possibilities of supporting relationships. He further indicates that although theoretical rationale leads us to assume that popularity is positively associated with self-esteem, conclusions from his study indicate that popularity is associated with behavioral poise rather than the individual's judgments of his own worthiness.

A major issue in the measurement of the concept of self lies in theoretical assumptions about the dimensions of the construct being measured. It shares the age-old problem of obtaining an objective measure of subjective experiences. In this study, it is assumed that the measure of self and the measure of others are primarily in the affective domain rather than in cognition. The rationale of this study is to identify the affective structures in the independent factors of the concept of self derived from a self-report instrument. In essence, the study attempts (1) to identify independent factors of self-concept based on subjective reports of a select sample of fifth and sixth grade children, (2) to describe the factorial dimensions of social relationships based on peer evaluation, and (3) to examine correlates of the dimensions of self with interpersonal constructs within the classroom.

Method

Subjects were 204 elementary pupils in grades 5 and 6 selected from two schools in two metropolitan areas in Georgia. One school had a population of more than 90 percent white, advantaged children, while the other school's population reflected more than 90 percent black, economically deprived (parental income below the poverty index) children. All students in these grades in both schools were included in the study resulting in approximately equal numbers of students from each school.

The Coopersmith SEI (3) was used to assess attitudes toward self while twelve adjective pairs in the semantic differential (SD) were used to rate several peer concepts in the classroom. Although Coopersmith stipulated five theoretical factors from his instrument it was decided to derive five factors of self-concept from the fifty-eight items by a factor analysis. Each S was instructed to complete the SEI by checking the appropriate box to indicate whether or not the statement described "how he feels about himself." Every statement was read by the examiner to eliminate inadequate responses attributable to reading ability. The team of four examiners administered both the SEI and the SD to Ss in both schools.

The SD consisted of twelve bipolar adjectives that the evaluator was able to rate on a 7-point scale. In order to obtain peer ratings of the children, each class was divided randomly into three groups and members of each subgroup completed the SD for each member of their own subgroup. Thus, it was possible to obtain approximately eight ratings of each child by his peers.

Results

The total number of responses for the 204 elementary school boys and girls on the twelve SD scales, was factor analyzed by the principal components solution with unities in the diagonal. Three factors satisfied Kaiser's criterion of roots greater than one, and the three factors were rotated by the varimax procedure.

Table 1.—Rotated Factor Loadings on the Semantic Differential for Peer Concepts (N = 855)*

Adjective Pairs	Factors			h^2
	Evaluation	Potency	Activity	
large-small		.77		.60
unpleasant-pleasant	.79			.67
fast-slow			-.69	.55
dull-sharp	.63		.39	.56
thin-thick		-.81		.66
happy-sad	-.67			.52
weak-strong		-.44	.60	.59
good-bad	-.81			.67
moving-still			.66	.44
unfair-fair	.75			.57
not active-active			.71	.56
heavy-light		.85		.73

*Loadings below .39 omitted

Factor loadings for the responses in rating their classmates on the semantic differential are found in Table 1. Factor I is easily identified as Evaluation based on significant loadings of evaluative adjectives: unpleasant-pleasant, happy-sad, good-bad, fair-unfair, and dull-sharp. Potency appears as Factor II (Table 1) with the highest loadings from adjectives describing power and ability: large-small, thin-thick, weak-strong, and heavy-light. Activity is well described in Factor III with the adjectives fast-slow, dull-sharp, weak-strong, moving-still, active-not active.

All responses on the SEI were factor analyzed. Intercorrelations were obtained among the fifty-eight scales and principal axis solution was obtained with unities in the diagonal. Based on Kaiser's criterion, five factors were extracted and rotated by the varimax procedure.

The major factor loadings for the factor structure of the SEI, are reported in Table 2. Although the factors are not as clearly identified as the constructs on the SD, five independent factors emerged, contributing 65 percent of the total cumulative variance. Two of these factors, I and III, reveal an essentially negative self-image, whereas factors II and V identify positive feelings about self. The fourth factor contains many of the items placed in the SEI to form the lie scale.

Factor I seems to be a relatively strong factor with significant loading on items describing how the child views himself. Items of self-rejection outweigh those of rejection by peers, parents, or teachers. This suggests that Factor I can be labeled the Self-Rejection factor within the total construct of the self concept. The eleven items contributing to Factor II seem to share some communality with parental concern and should be labeled, Parental Approval.

Factor III is probably a large general factor of negative self-esteem represented by highest loadings on items representing parental disapproval. In addition to the unhappiness in the home, a general self-rejection in school and social situations exists. Factor III was named Rejection by Authority.

Factor IV loads significantly on items that comprise the lie scale of the SEI. The inconsistency expressed by items 13 and 15, for example, appears to indicate that this factor is best called a Lie Scale.

Factor V appears to identify the independent, self-confident, poised child. It loads on items revealing peer, self, and adult approval. It is called the Social Acceptance factor of the self-concept.

To test the hypothesis that the semantic differential responses and the SEI responses belong in the same system of relationships, the three factor scores for each student on the SD and the five factor scores on the SEI were correlated by use of a canonical correlation.

Table 3 presents the one significant canonical correlation (R = .325) and its corresponding canonical vectors. These vectors are regression weights and indicate the mathematical strength of the factors in each subset in developing the relationship between factors of self-concept and peer concept. The vector weights indicate that two factors from the self-concept set are primary contributors while two other factors from this set contribute slightly less to the canonical correlation. The largest contribution to the canonical correlation of SD and SEI variables in this study

Table 2.—Factor Loadings for 212 Pupils on Items of Coopersmith SEI

Item No.	Items	I	II	III	IV	V	h_2
	FACTOR I: Self Rejection						
22.	I give in very easily	.31			.50		35
25.	I would rather play with children younger than I am	.52			.33		44
26.	My parents expect too much of me	.30			.36		39
31.	Things are all mixed up in my life	.45		.50			51
33.	No one pays much attention to me at home	.31		.52			46
37.	I don't really like being a boy–girl	.33			.35	.31	39
38.	I have a low opinion of myself	.53		.30			43
39.	I don't like to be with other people	.48			.33		46
42.	I often feel upset in school	.40		.44			40
43.	I often feel ashamed of myself	.54		.30			39
44.	I'm not as nice looking as most people	.51					30
46.	Kids pick on me very often	.46		.43			46
48.	I always tell the truth	.42					32
49.	My teacher makes me feel I'm not good enough	.59					40
50.	I don't care what happens to me	.60					42
51.	I'm a failure	.71					61
52.	I get upset easily when I'm scolded	.40		.44			42
53.	Most people are better liked than I am	.34		.52			40
54.	I usually feel as though my parents are pushing me	.48		.32		.33	47
56.	I often get discouraged in school	.53					37
58.	I can't be depended on	.60					51
	FACTOR II: Parental Approval						
4.	I'm easy to like		.42				27
5.	My parents and I have a lot of fun together		.52				34
10.	I can make up my mind without too much trouble		.34			.40	37
11.	I'm a lot of fun to be with		.60	.35			50
14.	I'm proud of my school work		.62				45
19.	My parents usually consider my feelings		.62				43
21.	I'm doing the best work that I can		.37	.41			35
24.	I'm pretty happy		.50				32
29.	I understand myself		.37			.39	35
36.	I can make up my mind and stick to it		.35			.55	43
47.	My parents understand me		.66			.32	57
	FACTOR III: Rejection by Authority						
3.	I often wish I were someone else			.42			34
7.	I find it very hard to talk in front of the class			.35			25
8.	I wish I were younger			.44			31
9.	There are lots of things about myself I'd change if I could			.47			33
12.	I get upset easily at home			.63			42
15.	Someone always has to tell me what to do			.40	.35		38
16.	It takes me a long time to get used to anything new			.46			36
17.	I'm often sorry for the things I do			.53			32
21.	I'm doing the best work that I can		.37	.41			35
27.	I like everyone I know			.33			26

(Continued)

Table 2.—Factor Loadings for 212 Pupils on Items of Coopersmith SEI—Continued

Item No.	Items	I	II	III	IV	V	h^2
30.	It's pretty tough to be me			.41			51
31.	Things are all mixed up in my life	.45		.50			46
33.	No one pays much attention to me at home	.31		.52			28
35.	I'm doing as well in school as I'd like to			.41			43
38.	I have a low opinion of myself	.53		.30			46
40.	There are many times when I'd like to leave home			.57			40
42.	I often feel upset in school	.40		.44			39
43.	I often feel ashamed of myself	.54		.30			46
46.	Kids pick on me very often	.46		.43			46
52.	I get upset easily when I'm scolded	.40		.44			42
53.	Most people are better liked than I am	.34		.52			40
54.	I usually feel as though my parents are pushing me	.48		.32			47
	FACTOR IV: Lie Scale						
6.	I never worry about anything				.49		36
11.	I'm a lot of fun to be with		.60		.36		50
13.	I always do the right thing				.69		58
15.	Someone always has to tell me what to do			.40	.35		38
20.	I'm never unhappy				.59		48
22.	I give in very easily	.31			.50		35
25.	I would rather play with children younger than I am	.52			.33		44
26.	My parents expect too much of me	.30			.36		39
34.	I never get scolded				.59		46
37.	I don't really like being a boy–girl	.33			.35	.31	39
39.	I don't like to be with other people	.48			.33		46
	FACTOR V: Social and Self-Acceptance						
2.	I'm pretty sure of myself					.40	28
10.	I can make up my mind without too much trouble		.34			.47	37
18.	I'm popular with kids my own age					.47	30
23.	I can usually take care of myself					.53	38
28.	I like to be called on in class					.52	36
29.	I understand myself		.37			.39	35
32.	Kids usually follow my ideas					.44	34
36.	I can make up my mind and stick to it		.35			.55	43
37.	I don't really like being a boy–girl	.33			.35	.31	39
41.	I'm never shy					.48	35
45.	If I have something to say, I usually say it					.65	45
47.	My parents understand me		.66			.32	57
54.	I usually feel as though my parents are pushing me	.48		.32		.33	47
55.	I always know what to say to people					.56	45
57.	Things usually don't bother me					.51	36

comes from Factor III, Activity, of the peer concept set.

Discussion

Based on inferences from the responses about their peers, elementary children in this sample perceive their peers in terms of three independent dimensions, which were labeled Evaluation, Potency, and Activity. As mentioned earlier, these findings are consistent with results from numerous studies using the SD on varied samples. The procedure for identifying all classroom children in semantic space not only reflects the many person concepts in a class but provides an opportunity to measure clusters of *concepts* and even the classroom climate. The measurement of the total feelings about significant others in the class offers a sociometric vehicle with which we can make further investigations into the concept of self.

The factors derived from the SEI indicate that the attitudinal dimensions of the self-concept do include self-perceptions as well as perceptions of social evaluation. Factor I appears to identify the child who has little confidence in himself, as the item loading highest on this factor states, "I'm a failure."

Factor II seems to identify the easy-going, happy, self-confident child who enjoys great parental support. It is interesting to speculate

Table 3.—Weights for the Vectors Corresponding to the Canonical Correlation (R=.325) Between Coopersmith SEI and SD Factor Scores

Self-dimensions	Factor	Weights
Self Concept:	I-Self Rejection	.656
	II-Parental Approval	.057
	III-Rejection by Authority	.713
	IV-Lie Scale	.505
	V-Social and Self Acceptance	.438
Peer Concept:	I-Evaluation	.360
	II-Potency	.359
	III-Activity	1.115

to express a high level of self-esteem as he reaches adolescence or whether self-esteem at that age may be less dependent upon parental acceptance.

Factor III appears to identify the child who experiences parental rejection and is marked by feelings of dependency, guilt, and confusion. One of the clearest messages revealed in this factor is that the child wants to be a different person. He appears to be less defensive and perhaps would be more susceptible to a change in behavior through therapy than the child identified by Factor I.

Factor IV appears to be primarily a lie scale. Interestingly, the desire to be of the opposite sex appears on this factor as well as on Factors I and V. It is possible that a sex bias may be involved in selection of this item.

The fifth factor identifies most clearly the child of high ego-strength, with the characteristics of a potential leader. He does feel that his parents are "pushing him" and expresses a desire to be of the opposite sex.

The significant canonical correlation suggests that Activity, as seen by peers, relates to self-concept factors. The uniqueness of this Activity factor receives much of its strength from the Self-Rejection and Rejection by Authority factors of the self-concept with less weight coming from the Lie Scale followed by Social and Self-Acceptance. The correspondence between these factors of peer and self-concept suggest that the child with high negative self-esteem is more apt to be included in the Activity factor, as viewed by his peers. This finding is not inconsistent with Coopersmith's thesis, referred to earlier, that the child with poise and self-assurance is more apt to be viewed favorably by his peers.

FOOTNOTE

1. We want to express our appreciation to the Research and Development division of the Atlanta Public School System for assisting us in locating the population of children.

REFERENCES

1. Brookover, W. T.; Patterson, A., "Self-Concept of Ability and School Achievement," *Sociology of Education,* 37:271–278, 1964.
2. Combs, C., "Perception of Self and Scholastic Underachievement in the Academically Capable," *Personnel and Guidance Journal,* 43:47–51, 1964.
3. Coopersmith, S., *The Antecedents of Self-Esteem,* W. H. Freeman and Company, San Francisco, California, 1967.
4. Gordon, I. J., *How to Study the Child in School,* John Wiley and Sons, New York, 1966.
5. Rentz, R. R.; White, W. F., "Congruence of the Dimensions of Self-As-Object and Self-As-Process," *The Journal of Psychology,* 277–285, 1967.
6. Wylie, R., *"The Self Concept,"* University of Nebraska Press, Lincoln, 1961.

FIELD NEWS

Willard E. North
Central Missouri State College

THE U. S. OFFICE OF EDUCATION is currently undertaking, by the funding of three different projects, a targeted program of research and development in reading. The three projects represent effort by the U. S. O. E. to provide an information base for both the funding of subsequent research in the field of reading and for a national literacy campaign, sometimes called the "Right to Read" campaign.

One of these projects, number 3, begins with the basic hypothesis that there exists a significant, qualifiable deficit between the present state of reading ability in the U. S. and the individual and social literacy needs of the populace. Operationally, project number 3 is an attempt to test this hypothesis through a critical examination of current literature in several areas of the field of reading. Specifically, it has the following three tasks: (1) To determine the extent and distribution of the national reading problem, by identifying, analyzing, and summarizing existing survey and test data which indicate the reading ability of various population groups found in the U. S. according to learner characteristics such as sex, age, race, IQ, etc., and environmental characteristics such as social economic status, geographic area, language spoken in the home, level of formal education obtained by the parents, and so forth. (2) To determine the use frequency and use distribution of instructional methods, approaches, procedures, materials, and equipment for reading instruction by pursuing the following questions:

> *(a) What methods, materials, approaches, equipment, and procedures are used to teach reading in the U. S., and to what extent are the major items in these categories used: i.e., how many of the learners are exposed to them?*
>
> *(b) What methods of reading instruction are built upon essentially different pools of basic knowledge?*
>
> *(c) How much time and resources are expended directly upon developmental reading instruction and remedial reading instruction?*
>
> *(d) What relationship between methods of reading instruction and reading achievement of the various subgroups of the population can be shown?*

(3) To describe the nature and extent of current practice in the training of those who teach reading by examining: the requirements of all 4 year institutions which train at least one hundred elementary teachers per year; the certification of standards for elementary school teachers, reading specialists, reading supervisors, in the fifty states and the District of Columbia; the relationship between teacher performance and teacher training.

The development of a bibliography relevant to the three problems of the project has been the task of the following committee: Sterl Artley, University of Missouri, Columbia; Director, University of Missouri Reading Resource Center in the ERIC/CRIER Network. Martin Carnoy, Stanford University, Palo Alto, California. John B. Carroll, Senior Research Psychologist, Educational Testing Service, Princeton, New Jersey. Martha Maxwell, University of California, Berkeley; Director of Reading Resource Center in the ERIC/CRIER Network. Constance McCullough, San Francisco State College, San Francisco, California.

The sources of documents have been varied, and a great deal of effort has been made to be as comprehensive as possible. They have included bibliographies suggested by each member of the advisory committee, ERIC, the Education Index, Psychological Abstracts, Sociological Abstracts, and other journals, the publications of Government agencies, sources in industry, state departments of education, selected educational research institutes and individuals, selected local school district research, foundations, the publishers of both reading materials and tests, and teacher training institutions.

The actual selection of bibliography involves a two-step process. First, all references are submitted to the members of the advisory committee. The members evaluate these references according to (a) the quality of their data as to representativeness and (b) their relevance to one or more of the specific research tasks. These evaluations are then combined, and a consensus as to inclusion is reached. Secondly, the references on which there was agreement are sorted according to the particular subtask to which they are relevant. If it becomes apparent that there are too many references in a single subtask, these are then resubmitted to the advisory committee to be given priority ratings. Further, if it becomes apparent that there are no references for certain tasks, the advisory committee is asked to help find references to fill these gaps.

The general policy for the review is that no reference is included in the bibliography without having been evaluated and ranked by all members of the advisory committee. Some references on which there was not total agreement as to worth or relevance will be included. Readers are invited to contribute relevant literature which might not be tapped by the research method outlined, for example, unpublished research papers. Comments and suggestions concerning the project are also invited. Correspondence should be directed to Educational Testing Service, 1947 Center Street, Berkeley, California 94704.

* * * * * *

NEW VOICES IN EDUCATION, a graduate student quarterly journal is a recent addition to educational publications. All articles are written by students. It is published by gradute students in education at the University of Florida. Students from over seventy-five other institutions participate in the production of this journal as associate editors and campus representatives. The journal is intended to provide "a form for a voiceless segment of the profession" and to furnish professors and administrators with current student thought and concerns.

The theme of the first issue was "Conflict in the Graduate Students." "Can and Should Public Education Survive?" and "Concepts of New Education" will be themes for the March and June issues respectively. Subscription costs are $3.00 per year for students and $5.00 per year for all others. Student manuscripts should be submitted to Content Editor, Box 2456 Norman Hall, University of Florida, Gainesville, Florida 32601.

* * * * * *

SOME 1,375 CHILDREN WHO ENTERED first grade in Bethlehem, Pennsylvania, in 1964 have been tested through the years to see if the two-thirds who learn to read with the Pitman Initial Teaching Alphabet (ITA) fare better scholastically than the one-third who started with the traditional ABC's. Of the children who still remained in the Bethlehem schools at the end of the sixth grade, a report shows that ITA can dramatically cut failure and that its effectiveness is felt long after the ITA code-breaking has been accomplished and the transition to regular alphabet made.

The study, financed initially by the Ford Foundation's Funds for the Advancement of Education and later also by the U. S. Office of Education, has been conducted by Dr. Albert J. Maxurkiewicz, Professor of Education, Newark (N.J.) State College.

Although the longitudinal study was weighted to favor the children in the traditional orthography (T.O.) control group (the ITA group included more of the highly mobile, less advantaged population), the ITA pupils showed superior performance on several counts: (1) the rate of failure as evidenced by the number of children who repeated grades, has been three times higher for the T.O. group (29.6% to 9.1%); (2) twice as many T.O. pupils (14.6% to 7.1%) received remedial reading as did ITA pupils; (3) on a subtest of a standardized battery, the ITA pupils scored higher on capitalization, knowledge in use of references, and dictated spelling.

One interesting discovery of the 6 years of testing, according to Dr. Maxurkiewicz, *"was that type of remedial reading needed by the 7.1 percent of the ITA pupils and 14.6 percent of the T.O. pupils differed. The T.O. pupils needed help on word recognition and comprehension, while the ITA child who had mastered the reading skills, needed additional help only in the comprehension area."*

* * * * * *

THE UNION FOR EXPERIMENTING Colleges and Universities has received a grant of $400,000 from the Ford Foundation for planning and development of its new *University Without Walls* program. This grant, along with a grant by the U. S. Office of Education, brings to $815,000 the amount available for *University Without Walls* planning and development.

A total of nineteen institutions will take part in the program: the University of Minnesota; the University of Massachusetts; Antioch College; New College at Sarasota; Shaw University; the University of South Carolina; Roger Williams College; Bard College; Chicago State College; Goddard College; Howard University; Friends World College; Northeastern Illinois State College; Stephens College; Loretto Heights College; Staten Island Community College; Skidmore College; Morgan State College, and New York State University.

The program will be opened to a broad age range of students—from 16 to 60 and older. It will provide highly individualized and flexible approaches to learning, making use of a wider array of resources for teaching and learning than is now the case. In addition to classroom work, key features of the program include the use of internships, field experience, and research assistantships as a regular part of the students' undergraduate curriculum, individual study, individual and group projects, seminars-in-the-field, video tape, and other media resources. There will be no fixed curriculum and no uniform time schedule for the award of the degree, with programs individually tailored and worked out between the student and his teacher–advisor. An *Inventory of Learning Resources* will be compiled to serve as a key guide for students and advisors in the planning of program sequences. An adjunct faculty of government officials, business executives, persons from community agencies, scientists, artists, writers, and other specialists will be used extensively, both in tutorial relationships and to conduct seminar-in-the-field programs.

Most institutions will begin their programs in the fall with about seventy-five students; some have already started with small pilot groups of fifteen to twenty students. Students participating in the *University Without Walls* at a particular institution will be able to take part in and make use of the resources of the other *University Without Walls* programs.

It is the intent of the program to develop new models and new forms for American higher education. Faith is placed in the student and his capacity for learning on his own, while providing close and continuing contact between student and teacher. The role of the teacher is redefined as a facilitator and co-participant in the planning and design of the student's learning experience. It is hoped that through the use of a new mix of age ranges that the program will encourage a new dialogue and increase trust between younger and older persons as they join together in a wide range of task-oriented and other kinds of educational experiences.

* * * * * *

INCREASED FEDERAL FUNDS will raise school spending for audio–visual instructional materials by $17.5 million to $90 million during 1971, according to studies made by Market Data Retrieval (MDR), a New York Education Market Research Firm. The source of additional funds will be a $37.5 million increase in money available to schools under Title II of the Elementary-Secondary Education Act which provides grants for school purchases of library and audio–visual materials. During 1971, a total of $80 million will be provided to schools for purchase of such materials under Title II, as compared to $42.5 million under the same act in 1970.

In addition to the $30 million in Title II funds, elementary-secondary schools will spend $60 million in funds from state, local, and other Federal government sources for audio-visual instructional materials, according to a report from this firm.

Studies by MDR indicate that growth in audio-visual spending would be level or show a slight decline this year were it not for this infusion of federal funds. Herbert M. Lobsenz, president of MDR, believes that while these funds will increase the amount of audio-visual equipment purchased in 1971, the long-term effect will be less beneficial. A pattern may be started in 1971 that will not be sustained in 1972. Secondly, he feels that this pattern may deter normal acquisition of equipment by persuading school systems to await periodical infusions of Federal funds to build up their audio-visual and library collections.

A state-by-state analysis of 1970–1971 Federal funds for education is available from Market Data Retrieval, 22 West 48th Street, New York, New York 10036. The price is $17.50.

* * * * * *

In 1970 THE EDUCATIONAL RESEARCH Association of New York State underwent a transition to become the Northeastern Educational Research Association (NERA). Abstracts of the proceedings for the tenth and eleventh convocations (1969–1970) are now available from the Association. These abstracts are published in two volumes with about sixty research abstracts in each volume.

Copies may be obtained from: Dr. Richard M. Clark, Department of Educational Psychology, State University of New York, 1400 Washington Avenue, Albany, New York 12203. The cost is $2.00 per volume or $3.00 for both volumes.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

GORDON S. BIGELOW is Dean of Student Group Affairs and Assistant Professor of Counseling and Personnel Services at Kent State University. Prior to going to Kent he spent a post-doctoral year in counseling psychology at the University of California, Berkeley. He has taught as a Counselor Educator for the past several summers at the University of Maine, and had several years of experience as a high school counselor and as a social case worker. He holds an AB in Sociology from Brown University, an MEd in Counseling and Guidance from the University of Maine, and an MA and PhD in Educational Psychology from BYU where he was an NDEA Graduate Fellow.

CLAUDE W. GRANT is Professor and Chairman of the Department of Educational Psychology at the University of Utah. He has been Director of Institutional Studies at the University of Utah and has held positions with New York University, Syracuse University, and the University of Minnesota. He has been an active consultant to public school systems and to the psychology training program in the Veterans Administration. His BS is from the University of Houston, MS from the University of Utah, and PhD from the University of Minnesota.

KARL E. KEEFER holds the MEd degree in Educational Administration from the University of Chattanooga and the EdD in Educational Psychology and Guidance from the University of Tennessee at Knoxville. His professional career has involved him in administration and teaching in Tennessee, Washington (D. C.), and South Carolina. After serving as Associate Professor of Educational Psychology at the University of Tennessee—Memphis, State University Center for Advanced Graduate Study in Education, for several years, he became Professor of Education and Dean of the School of Education at the University of Tennessee at Martin in the fall of 1969, where, in addition to his administrative responsibilities, he teaches graduate courses and engages in field service activities to school systems in west Tennessee.

WILLIAM A. POPPEN and CHARLES L. THOMPSON are Assistant Professors of Educational Psychology and Guidance at The University of Tennessee in Knoxville. Dr. Poppen received his BS from South Dakota State University and his MA and PhD from The Ohio State University. His interests are in the areas of elementary school counseling and interdisciplinary approaches to school guidance. Dr. Thompson received his BA and MA degrees from The University of Tennessee and his PhD from The Ohio State University. His writings and teaching have focused upon the counseling relationship and behavior change in children.

BERT O. RICHMOND is an Associate Professor in the school psychology program at the University of Georgia. His major research efforts are directed toward enhancing educational experiences of the culturally disadvantaged. WILLIAM F. WHITE is an Associate Professor at the University of Georgia where he has been since 1964. He has reported more than seventy-five studies in research journals and has published three books. Research interests presently are focused on psychosocial variables in the classroom, especially among economically deprived.

RICHARD F. SEBOK is Director of the Counseling Center and a Departmental Instructor in Education at the University of Texas at El Paso. He obtained a BA from Millikin University, Decatur, Illinois, and a master's from the University of Illinois, Urbana. He previously served as a junior high principal, an Associate Director of Institutional Studies, and an Instructor of Educational Psychology. He has held his present position since 1969.

SUSHILA SINGHAL served as a consultant to the New York State Education Department in the Bureau of Pupil Testing and Advisory Services at the time she worked on the study of migrant gains. She obtained her BA and MA from Agra University and her PhD from Calcutta University. She has now returned to India. PRISCILLA H. CRAGO, prior to her retirement, was Chief of the Bureau of Pupil Testing. She holds an AB from Swarthmore and an EdM from Rutgers. She is a member of Kappa Delta Pi. Both authors have published a number of papers on testing in professional journals.

LINDLEY J. STILES is Professor of Education for Interdisciplinary Studies, Sociology, and Political Science at Northwestern University. He was previously Dean, School of Education, The University of Wisconsin and Co-director of the Research and Development Center for Learning and Re-education at The University of Wisconsin. His experience in education ranges from being an English teacher in junior and senior high schools to Director of Graduate Studies in Education. He holds an AB, MA, and EdD all from the University of Colorado. Besides his degree work he has studied teacher education and education programs in key nations around the world. He has published over 150 professional articles and is a regular columnist to The JER.

ROBERT WILLIAMS, S. M., a social studies instructor at Chaminade High School, Hollywood, Florida, received his BSEd from the University of Dayton. He is currently working toward an MS at Dayton in an educational research training program funded through the Southwestern Ohio Educational Research Council by the U. S. Office of Education, Bureau of Research. His interest has been to relate to elementary teachers the value of Piaget's theories of the sequential developmental stages of the child's thought in relation to specific subject areas.

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor of Curriculum and Instruction, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Associate Professor of Special Education, Yeshiva University, New York.

GEORGE S. JURATA, Coordinator of Reading, Jefferson County Schools, Colorado.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

S. JAY SAMUELS, Associate Professor of Educational Psychology, University of Minnesota, Minneapolis.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Editors Emeriti

HARRY J. BAKER, 2241 Q Via Puerta, Laguna Hills, California 92653.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

JeR

51st Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 64 JULY-AUGUST 1971 NUMBER 10

ARTICLES

Curriculum Development and Evaluation—*Hulda Grobman* 435

Recency of Secondary School Mathematical Preparation as an Influence on First-Year College Mathematics Achievement —*Alton R. Goddard* 443

The Interrelationships Among Self-Esteem, Personal Values, and Interpersonal Values—*Sandra Lou Stein* 448

A Longitudinal Study of Visual Perceptual Training and Reading Achievement—*Neal R. Gamsky; Faye Williams Lloyd* 451

Effects of Same Sex Class Organization on Junior High School Students' Academic Achievement, Self-Discipline, Self-Concept, Sex Role Identification, and Attitude Toward School—*Joseph R. Ellis; Joan L. Peterson* 455

Nature and Language of Scientific Terms in the Arabic Press of the Middle East—*Wadi' D. Haddad* 465

Academic Performances of Matched Groups of Community College and 4-Year College Students—*Lincoln H. Hall* 475

The Effect of Selected Media Feedback Upon the Interactive Behavior of Student Teachers—*B. Charles Leonard; Frederick John Gies; Jon S. Paden* 478

Yearly Index 484

Departments

Among Those Present 483

Book Reviews—*Robert E. Clasen* 454

Field News—*Willard E. North* 481

Forthcoming Articles 450

It Says in The JER 434

New Books from the Publishers 447

Research Briefs—*Daniel S. Parkinson* 442

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin. Copyright, 1971 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, Supervisor of Editorial Services.

It Says in The JER...

CAN SUCCESS LIMIT A PROJECT? Grobman feels this may be what is happening in curriculum development, as more groups become interested, funds have to be divided into smaller parts thus limiting each particular program (p. 435).

SCORES ON SAT (MATH) and CEEB (Math) along with achievement in first-year college math for students who had taken a math course their last year in high school and those who had not were compared. Achievement was greater for those who had had math their last year in high school. Goddard based his comparison on an analysis of covariance procedures (p. 443).

SELF-ESTEEM AND the personal value of achievement for high school seniors was the only statistically significant correlation between self-esteem and six personal values studied by Stein. She used the Coopersmith Self-Esteem Inventory and Gordon's Surveys of Personal Values and Interpersonal Values (p. 448).

RURAL SCHOOL DISTRICT kindergarten students participated in the Frostig Program for the Development of Visual Perception. Gamsky and Lloyd found indications that the program benefitted the students' reading ability when they were tested on the Stanford Achievement Test the following year (p. 451).

JUNIOR HIGH SCHOOL GIRLS are more against sex-segregated classes than are the boys. Ellis and Peterson found that girls also tended to get higher marks from teachers, score higher on achievement tests, and have more favorable attitudes toward social groups (p. 455).

IS THERE A CONSISTENCY in the usage of scientific terms in the Middle East press? Haddad found that reading of scientific articles required substantial technical knowledge as the terms used varied from being totally or partially borrowed. There was also a problem of translation and transliteration leading to inconsistencies in usage of terms (p. 465).

DO STUDENTS BEGINNING college at a junior college get higher mean GPA's than students starting at 4-year institutions? Hall found that they earn higher grades during the first 2 years but there is no significant difference after transfer to a 4-year institution (p. 475).

DOES SELECTED MEDIA feedback affect the interactive behavior of student teachers? Leonard, Gies, and Paden found that those receiving video feedback utilize less direct teaching influence while there was no statistically significant change within the audio group (p. 478).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Editorial/Abstract

The State of the Art:
Curriculum Development and Evaluation

◆ Two new factors in the last decade, massive public funding and a developmental approach to curriculum building, have brought major changes in curriculum and the development of a new field of research—curriculum evaluation. A high degree of conformity existed across the United States in subjects taught at particular grade levels and in textbooks in use. The National Science Foundation in 1957 began funding developmental approaches to curriculum innovation in science and mathematics. Federal financing of teacher training programs gave further impetus for change in the schools. Further, the change has occurred at all levels of community education.

Several concepts are related to these developments: the ideas of inquiry and the structure of the discipline; the taxonomy of educational objectives and behaviorally stated objectives.

One deficiency in recent developments has been articulation between grade levels and between subjects. In part this is due to speed of developments and difficulty of integrating fields being simultaneously developed.

Evaluation developed to provide feedback on preliminary material in use and to describe the finished product. A series of new concepts has been put to use in relation to the use of developmental approaches: antecedents, transactions, outcomes, formative evaluation, summative evaluation, evaluation of process and of product, micro-evaluation, and macro-evaluation.

This analysis is very helpful in understanding curriculum developments of the last fifteen years and leaves the reader with optimism based on the positive changes and pessimism based on the current slowdown due to the competition for limited funds. However, the success with which change has been brought about in the schools is encouraging as one looks ahead.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Curriculum Development and Evaluation

HULDA GROBMAN
New York University, Washington Square

IN THE LAST decade, the whole notion of how curriculum is built, of what constitutes curriculum, and of the parameters of curriculum research have undergone drastic change. Instead of talking about the need for change, major change has been occurring in subjects taught, content within subjects, grade level placement, and approach to the subject, and a new field of curriculum-related research—curriculum evaluation—has emerged. Perhaps this increased rate of change is most directly attributable to two new factors: massive public funding for curriculum preparation and a new method of preparing curriculum.[1]

Until about 10 years ago, despite disclaimers to the contrary, there was a high degree of uniformity in the content of courses taught throughout the United States, a situation which had existed for many years. Not only were the same subjects taught at a given level, but often only one or two quite similar textbooks were in general use as the major instructional tool for a subject at that grade level. From one decade to another, changes were largely in the addition of new facts, rather than in basic approach and treatment.

Basic change was not occurring, despite periodic investigations of education by national educational and scholarly groups in various disciplines, and despite their strong recommendations for change (15). No single author or publisher could afford to undertake the responsibility and the financial risk of producing radically new materials; if he had, such materials would probably have been unused. The system for training teachers was geared to existing materials rather than to change; college admissions and college admission tests were geared to standard courses and to standard course coverage; and there was a general conservative tenor in most school systems. The entire educational system seemed designed to prevent change.

Curriculum Development

Throughout the 1950's there was a foment of discontent with public education, which reached a climax with the Russian launching of Sputnik, the world's first space probe. Many critics focused on the shortcomings of science and mathematics education in the schools as the cause of America's lagging in the space race. Concurrently there was an interest in a new tactic for producing student materials in science that would permit faster updating of curricula—*a developmental approach.*

Basically, this approach involves the production of new curricula, using experimental tryouts of preliminary materials and collecting feedback from such tryouts to be used for the improvement of curriculum prior to its release for general distribution. Most often, such efforts are group efforts, as contrasted with individual or co-author efforts. Thus, the *developmental curriculum project* or *curriculum project* refers to a group effort in producing curricular materials through a developmental process. In these cooperative efforts, frequently both educators and scholars in the subject areas are involved. The University of Illinois Committee on School Mathematics (UICSM) which initiated its work in 1951, may have been the first group to use a developmental process in curriculum change. However, large-scale use of this approach did not occur until the late 1950's when federal monies—channeled through the National Science Foundation (NSF)—became available for the first time for the production of high school materials in science and mathematics.

Although the idea of federal financing for curriculum work in physics preceded Sputnik, undoubtedly that event stimulated further action and increased the amount of funding available. Thus, in fall 1957, the Physical Science Study Committee (PSSC) received its first NSF grant to design a radically new high school physics course to better prepare students for college physics. (It is not clear whether the UICSM activities influenced the PSSC approach, or whether the PSSC approach, though in many ways similar, was developed independently.) PSSC is a landmark in that it was the first large-scale, federally financed developmental curriculum project. Never before had such sums (over $5 million in the first several years of PSSC) been spent to prepare curriculum materials for a course in a single subject area.

Era of Innovation

This opened a new era of curriculum innovation. Closely following were the School Mathematics Study Group (SMSG), Chemical Education Materials Study (CHEM Study), Chemical Bond Approach (CBA), and the Biological Sciences Curriculum Study (BSCS), all of them financed through NSF funds and using developmental processes in preparing curricula. Initially, NSF funding was limited to science and mathematics education at the senior high school level, but later public financing through NSF and the U.S. Office of Education (OE) was expanded to other levels of public school as well as preschool and college, and to virtually all subject areas. Private foundations have also been funding similar activities and, most recently, some commercial textbook publishers have supported their own developmental curriculum projects. The history of individual curriculum projects and descriptions of materials and approach to subject areas are included in A. Grobman (12), Wooton (30), McCoy (20), Goodlad (10), Goodlad with others (11), Merrill and Ridgway (21), and Lockard (18). H. Grobman (14) includes a comparative analysis of the projects in terms of decisions and processes in preparing curricula and in diffusing materials.

Those developmental projects in existence for several years have had an unprecedented success in effecting change in American education. While some schools in America remain unchanged, and not all the goals of the curriculum designers have been realized, real change has occurred. Science has become increasingly laboratory oriented. The term, "new math," has become part of the American culture. The focus on history is broadening to a focus on history in relation to other social sciences, and history as more than a collection of facts and dates. Astronomy, anthropology, sociology are appearing not only in the senior high school but are taught in an academically respectable fashion in elementary grades. Indicative of the effectiveness of the developmental method is the speed with which projects have multiplied in the United States and abroad. This would seem to indicate that this method of preparing curriculum will occupy a major place in American education for a long time, limited only by the availability of financing.

The practicality of effecting major change in subjects taught and in subject matter content within existing subject areas has been enhanced by the parallel widespread development of federally financed teacher training programs, initially established in science and mathematics but later extended to other fields. Such programs predated the developmental curriculum movement but reflected the same public concern with science and mathematics education. While these programs to upgrade teachers were generally initiated independently of the new curricula, throughout the 1960's many of the institutes and workshops established under these programs have been designed specifically to prepare teachers for use of the new materials. This has drastically reduced the delimitations on restructuring curriculum posed by the teacher's existing skills.

Other barriers to change—such as college admission requirements and examinations—have been diminishing, some more slowly than others, but change is clearly in sight. A major barrier to change in secondary curricula is the threat of the College Entrance Examination Board (CEEB) tests which have become more widely used in college admissions as the pressure of increased applicants has made the question of admission of students more difficult for colleges. Though change in the CEEB procedures has been discouragingly slow to date, portents for change are suggested in the recent report of the CEEB Commission on Tests (4).

Thus, much of the curriculum change in recent years is an outgrowth of the developmental process; much has been publicly funded; and the funding has been at an unprecedented level. Initially the change momentum focused on senior high school, first in science and mathematics and then in other areas. The next major focus was on elementary curricula. Junior high school curricula became a focus in the latter part of the 1960's. It should be noted that this sequence reflects the availability of federal funding for this purpose, and not the considered judgment of educators and subject-matter scholars as to relative priorities. Unfortunately the tardy focus on the junior high school years coincided with greater competition for funding, not only among the various educational priorities—with education of the disadvantaged receiving new levels of attention—but also with defense spending. Thus, curriculum reform for the junior high years has not kept pace, and adequate funding for this work is increasingly difficult to obtain.

Several commonalities or recurring themes are present in much of the developmental curricula. One is the participation of both educators and subject-matter specialists. This participation varies in degree of participation of the respective groups and the development phase at which participation takes place, but for the most part, both kinds of specialists participate.

Two common threads of the new curricula are the idea of *inquiry,* reflecting a learning by doing and by questioning, and the idea of *structure of the discipline.* While undoubtedly Bruner's *Process of Education* (4) was influential in popularizing these ideas, both were, in fact, integral parts of several curriculum projects prior to the Woods Hole Conference of which Bruner's volume was an outgrowth.

Another common referent in the development of these curricula is the *Taxonomy of Educational Objectives* (2, 16), reflecting a concern with different levels of cognitive skills and with the development of attitudes and values, though the former have been given more attention than the latter. (The psychomotor domain (26)—the area of manipulative skills—though often prerequisite to or concomitant of the desired cognitive outcomes, has received relatively little attention.)

Also influential to the development of many of the new curricula, particularly those starting later than the initial thrust, is the idea of behaviorally stated objectives, with Mager (19) as the most frequently cited source, though the idea of specifying objectives clearly was enunciated in the 1930's during the eight-year study under Ralph Tyler's direction and was underscored by Tyler in subsequent years (28). Unfortunately, one aspect that Tyler emphasized has been generally neglected in the preoccupation with behaviorally stated goals, that is, the determination of the appropriateness of the objectives. Thus, there has been more emphasis on precision in statement of objectives than on their appropriateness in terms of the broader tasks or long-range goals to be achieved. For some, the whole is assumed to be the sum of the parts; accomplishment of specific short-term objectives is assumed to contribute to the longer-range objectives. And few have systematically considered the interrelationship of affective, psychomotor, and cognitive domains and the relationship between immediate goals (often the only ones stated behaviorally) and the intermediate and long-range goals.

Increasingly, curriculum is being more broadly conceived by its developers as including all that is required to implement the basic purpose of the course. To a greater extent, curriculum materials go beyond the student textbook, laboratory manual, and workbook. Thus, *educational packages* are appearing frequently. These vary in extent to which all needed materials for teaching a given curriculum are included, but some extend beyond the customary student basic materials and teacher guides to include such items as the equipment and supplies needed to implement the course, various types of films, programmed materials, and readings integral to or supplemental to the curriculum, periodic written and laboratory tests, and teacher training materials including even materials to train those responsible for training teachers. This does not imply that all developmental curricula are all-inclusive packages. Some curriculum projects work on student text materials; some work on teacher materials in addition to or in place of student materials; some are concerned with supplemental learning materials; some are primarily process oriented to the extent that no student materials are prepared, but rather the emphasis is on preparing the teacher to operate in certain ways.

Individualization

The improved instrumentation of computers for programmed instruction and for computer-assisted instruction (CAI) has had a major influence in some of the developmental curricula, and has facilitated a focus on individualized instruction. Here there are two interpretations of individualization. In one, all students go through the same materials (either in straight-line or branched programming), but the pace varies. The other, perhaps best exemplified in Individually Prescribed Instruction (IPI), has available multiple pathways for reaching a stipulated goal; this is not merely branched programming but incorporates a wider concept of alternative routes, sequences, and choices.

The latter method is consistent with the concept of *mastery learning,* that is, the theory that virtually all (90%) students can achieve mastery of given objectives with the differences among students reflecting not the extent of achievement of goals but rather the time required and the means appropriate for achieving this mastery (1, 2). An additional essential ingredient is the provision of frequent non-punitive feedback on performance, so that the student is informed of how he is progressing and what his learning problems are. Also the student is insured that his ultimate reward (e.g., a grade) will not be prejudiced by problems encountered during the course of learning or by the pace of learning. Thus, all students can receive a grade comparable to A; it is hypothesized that this will enhance motivation as well as self-concept. Given this theory, a student's failure to learn reflects adversely on the curriculum and instruction rather than on the student. While the developmental process is highly appropriate for developing curriculum materials for mastery learning, the full implications of this theory for curriculum development are not yet clear. However, it may well provide the most exciting influence on curriculum development of the 1970's.

Most curriculum innovation of the 1960's has not focused on the questions of articulation among grade levels and among subject areas. The new curricula for a given subject area are generally prepared in isolation from other subjects taught at that grade level or at previous grade levels. Thus, science is still taught as discrete from social science; mathematics is discrete from science and social studies. Students who have studied—though perhaps not really learned—graphing in mathematics, express indignation when asked to interpret graphs in biology tests when no biology unit has focused on graph interpretation. Students who have been taught in social studies that inquiry is the most appropriate method for approaching problems may find this approach frowned on in their English classes.

In part, this situation reflects time pressure. Before one can coordinate courses, one must have the basic structure of feasible courses. Until quite recently there have not been enough inquiry-oriented courses available at the various grade levels and disciplines to plan a systematic coordination. Another delimiting factor is the textbook market. A 1-year textbook that can stand alone has a larger market than one that is part of a tight 3-year sequence. Thus, curriculum developers face the problem of whether to produce optimal materials that are relatively unused or to produce good, though less-than-optimal materials that have a potential for rapid adoption. A further difficulty is student mobility. If materials are built into tight sequences, with each unit building on previous learnings in that and other subjects, how will the student fare who transfers into or out of the system? This is a question on which little systematic investigation has been directed. At the present time, the problem of coordination of curricula among subject areas and articulation within subject area between grade levels while still retaining flexibility within the school systems remains one of the serious delimitations on effective education.

Curriculum Evaluation

The most immediate evaluation need of the developmental curriculum process was the feedback of information on preliminary materials during the process of materials preparation, to determine the general feasibility and to obtain clues concerning needed changes. This was particularly important in situations where radically new approaches to the subject matter were being initiated, where more sophisticated approaches were being attempted, where subject matter was being introduced at a lower grade level or being introduced for the first time at any grade level. For example, until someone tried to teach sets or number theory to grade school children, no one could be sure this could be done. It would be futile to develop and issue an entire curriculum based on these ideas unless one could be reasonably confident that they were feasible. Thus, the developmental process developed materials, tried them at a preliminary stage, sometimes in small segments, to determine feasibility, before continuing materials development. Clearly the kind of evaluation task implied here differs from that in conventional research investigations.

A second urgent evaluation need was to describe a finished product, to compare it with parallel products where previous curricula in the same subject existed or simply to describe it where no parallel curriculum was available for comparison. Less obvious evaluation needs related to the developmental process involved the assessment of various portions of a curriculum, the process by which the curriculum was prepared (including the sociology of the organization preparing the materials), the process by which schools adopt or do not adopt curricula, the way in which curricula are implemented once they are adopted.

Thus, the rapid development and adoption of new curricula in the 1960's was paralleled by a new kind of educational research focus—curriculum evaluation. That curriculum evaluation is a rather recent development is illustrated by the fact that the 1969 edition of the *Encyclopedia of Educational Research* (7) was the first to include an article on evaluation, and the *Review of Educational Research* did not focus on this area until the late 1960's. The dearth of pertinent literature and techniques to meet an emerging need led to the inauguration of a special publication series on curriculum evaluation by the American Educational Research Association (AERA) (23). This does not mean that there had been no earlier research on curriculum, but rather that for those curriculum investigations termed *evaluation*, the focus, the ground rules, the purposes, and the audiences are often different. These differences imply changes in techniques used, in research skills required, and in disciplines involved. Where earlier curriculum research had generally focused on psychological testing, using standardized instruments, a more appropriate current definition suggests that curriculum evaluation "*is the process of delineating, obtaining, and providing useful information for judging alternatives*" (27). Given this context, evaluation is concerned not only with learning outcomes resulting from use of materials, but starts prior to materials preparation and extends through preparation, dissemination, use, and re-adoption or discontinuance of use. The focus is on practicality—on providing a base for any decision making related to development and use of the curriculum.

There is no set boundary between research and evaluation or between evaluation of curriculum and evaluation of learning or of instruction. Some curriculum evaluation is research in the traditional sense; some is not. Some curriculum evaluation is evaluation of instruction or of learning; some is not. While the distinctions between research and evaluation and the nature of the parameters of evaluation had been touched on earlier (6, 13, 25), they are most clearly delineated in Tyler (29) and Stufflebeam and others (27).

Evaluation Methods and Problems

A number of guideposts for curriculum evaluation were established by Cronbach (6) and Stake (24) which added not only to the idea of what evaluation is and is not, but also have contributed to the new vocabulary which has helped to clarify the function and scope of curriculum evaluation. Thus, for evaluating radically new curricula, Cronbach justified failure to use conventional control groups and existing standardized tests; while these ideas are not radical today, in 1962 they were highly controversial.

Cronbach also distinguished between evaluation of curriculum and evaluation of individual learning. Stake added to evaluation vocabulary the terms *antecedents, transactions,* and *outcomes,* to identify the factors that must be accounted for in evaluation of student performance with a curriculum. Scriven (22) contributed the terms *formative evaluation* and *summative evaluation* to distinguish between evaluation during the development of materials and evaluation of the finished product. *Evaluation of process* has been used to refer to evaluation of the process of preparing materials as well as to the process of their adoption and implementation in the classroom, as distinguished from *evaluation of product,* which is concerned with the materials themselves and the outcomes resulting from their use. *Micro-evaluation* refers to the examination of small segments of a curriculum or of a particular variable, and *macro-evaluation* refers to broader questions such as the general suitability of the materials (13).

One of the most troubling aspects of curriculum evaluation has concerned selection of criterion measures. Educational research traditionally has depended on standardized achievement and attitude tests and rigorous criteria, for validity of such tests is available. Curriculum evaluation is broader than testing students. And even when classroom tests are used, these may not meet conventional test standards. For example, when a curriculum developer is preparing materials in a subject not taught before or not taught at a given grade level before, or when the emphases or desired outcomes differ from those of existing curricula, standardized tests do not exist; pretested instruments are generally not available; and when they are, preexisting norms may be inappropriate for the new target group.

Even the rules for constructing tests for developmental curricula may differ. For example, when a curriculum focuses on mastery learning, the usual test-construction protocols—that scores should average 50 percent correct, and that items should be eliminated if almost all students get them correct—are clearly inappropriate. These protocols are equally inappropriate when the purpose of the test is to find out how much students learned rather than to distribute students along a curve. Thus, the developmental curriculum quite appropriately may include in a test an item that virtually all students get correct, if the project is interested in knowing whether all students have the skill tested. The project may also wish to include some relatively easy items in the test simply to encourage students, since for project purposes it is far more important to have students who are not turned off by the test than to have a test which meets standards developed for other purposes.

One of the less fully discussed aspects of evaluation is the threat that seems implicit in evaluation activities (13, 14, 27). The various audiences for evaluation are generally apprehensive about the evaluation and the evaluator, and the very presence of the evaluator may constitute a threat. The perception of the evaluator often is of one who will judge, and the thought of being judged is not a happy one. This problem may result from a lack of understanding of how evaluation results can be used to strengthen programs. But perhaps it reflects a more basic difficulty—a lack of understanding of the whole idea of experimentation. Thus, the school board member, funder, or even curriculum project director who expects hundreds of experiments before a medical breakthrough on disease occurs may be completely intolerant of an educational experiment that does not produce optimal results the first time. There is an inadequate understanding among the various audiences for evaluation of the complexity of education, of the multitude of variables involved, of the nature of the human organism and its degree of amenability to change. Thus, evaluation is far more political than other educational research; it is far more ego-involving for the evaluator and for those involved in the program being evaluated, and apprehension concerning an evaluation may effectively sabotage the best of evaluation designs.

Ten years ago, curriculum evaluation was concerned with formative evaluation to provide immediate feedback to developmental curriculum projects to provide guidance to the curriculum developers. This was generally in rather gross terms, and included informal classroom observations, teacher reports of individual experience, teacher group discussions, classroom testing on specially designed classroom tests, and reviews by experts (educators, psychologists, and subject matter specialists). Summative evaluation, after the materials were completed, generally focused on comparisons among curricula, on rather gross measures of subject matter achievement and general attitude. These were of necessity rather crude devices, but there were no precedents to follow, few measuring instruments available, and enormous time pressures.

More recently, as more experience with evaluation has been garnered, as expertise is becoming available, as the appropriate scope of evaluation is more closely perceived, as time pressure is reduced, and as persons from a wider variety of specialties are becoming interested in evaluation, the focus is shifting to such questions as: How does change in school curriculum occur? What happens to materials after they are adopted? To what extent are they implemented in the classroom? Which parts of the materials are actually used? Is the method of use appropriate to the intent of the curriculum? For which students is the curriculum most appropriate? For which teachers or teaching styles is the curriculum most appropriate? Is differential sequencing of subsections in the curriculum a significant variable in student achievement? What other significant variables in the teaching/learning situation are relevant to the use of the curriculum? Are the curriculum objectives appropriate; that is, is the curriculum valid in terms of what it has set out to do? What are some of the unanticipated outcomes—desirable and undesirable—related to use of this curriculum?

As illustrated by Gallagher and others (9) and Lindvall and Cox (17), the focus of evaluation is increasingly on more precise micro- and macroevaluation. Not only are the measuring instruments more precise and their use more systematic, but the data collected are being examined more thoroughly.

Reports on aspects of curriculum evaluation as presently practiced appear in psychological and educational journals, but they also appear in journals in other fields as well. For this reason it is difficult for investigators to locate all pertinent references. However, useful recent bibliographies concerning the parameters and conventions of the field as well as of research in the field are found in the *Review of Educational Research* (8), Stufflebeam and others (27), and Bloom and others (3). The most comprehensive discussion of the present status of curriculum evaluation and its potential, its parameters, problems, and techniques is in Stufflebeam and others (27). Because of the stature of the authors, and also because this is the first comprehensive overview of the field, this volume may well become a definitive point of reference concerning the nature of evaluation, as well as a point of departure for later investigators in this area.

Conclusion

The future of the developmental method of preparing curricula seemed well assured 4 or 5 years ago. The earliest projects had released materials for general use. Many of these were extremely well received by educators and subject matter specialists, and met with rapid and widespread adoption. A way of rapidly changing curriculum along desired lines had been developed and proven. It was generally agreed that the process needed refinement, and that not all attempts were equally successful. Nonetheless, use of the method was spreading at an astonishing rate, as of 1970, in science and mathematics alone, there were some 150 developmental curriculum projects operating in the United States (18).

Perhaps it was the extent of this success that is a serious delimiting factor at the present time. Competing demands on public funds by other educational and non-educational programs set constraints on funding available for curriculum development. And, as more and more groups seek a share of the funds available for curriculum development, the total funds have been divided into smaller and smaller grants and stricter controls are being imposed on their expenditure. The seriousness of the situation may be indicated by the statement of the director of one of the early curriculum projects that, given today's grants and control procedures, his project could not have achieved success (12).

While commercial textbook publishers have sponsored some developmental curriculum work, their efforts cannot be as bold nor as extensive as those that are publicly funded. Further, for publisher's projects, the evaluation efforts in particular are under-supported, and for some, evaluation is nonexistent. Doubtless this reflects the fact that purchasers of textbooks do not ask for information on the formative and summative evaluation before adopting textbooks.

Curriculum evaluation is also suffering from budgetary constraints as federal funds for this purpose are drying up. Evaluation is expensive and some outside funding is generally necessary. Much of the work cannot be done by an independent investigator on his own time without outside resources. Further, the direction of curriculum evaluation may be changing radically as a reflection of the present preoccupation in education with cost/benefit ratios and with accountability. The kinds of questions that will be asked and the kinds of information that will be sought may come to reflect considerations of efficiency insofar as these are indicated by rather short-term and readily measurable outcomes rather than by more basic considerations.

FOOTNOTE

1. Some of the materials in this and the following paragraphs are adapted from H. Grobman (14).

REFERENCES

1. Block, James H. (ed.), *Mastery Learning, Theory and Practice*, Holt, Rinehart and Winston, New York, 1971.
2. Bloom, Benjamin S. (ed.), *Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I, The Cognitive Domain*, David McKay, New York, 1956.
3. Bloom, Benjamin S.; Hastings, J. Thomas; Madaus, George S., *Handbook on Formative and Summative Evaluation of Student Learning*, McGraw Hill, New York, 1971.

4. Bruner, Jerome S., *The Process of Education*, Harvard University, Cambridge, Massachusetts, 1960.
5. College Entrance Examination Board Commission on Tests, *Report of the Commission on Tests: I. Righting the Balance; II. Briefs, College Entrance*, Examination Board, New York, 1970.
6. Cronbach, Lee J., "Course Improvement through Evaluation," *Teachers College Record*, 64: (no. 8) 672–83, May 1963.
7. Ebel, Robert L. (ed.), *Encyclopedia of Educational Research*, Fourth edition, Macmillan Company, Collier–Macmillan Limited, London, 1969.
8. "Educational Evaluation," (issue title), *Review of Educational Research*, 40: (no. 2), April 1970.
9. Gallagher, James J.; Nuthall, Graham A.; Rosenshine, Barak, *Classroom Observation*, AERA Monograph Series on Curriculum Evaluation, No. 6, Rand McNally, Chicago, Illinois, 1970.
10. Goodlad, John, *School Curriculum Reform in the United States*, Fund for the Advancement of Education, New York, 1964.
11. Goodlad, John, with Von Stoephasius, Renata; Klein, Frances, *The Changing School Curriculum*, Fund for the Advancement of Education, New York, 1966.
12. Grobman, Arnold B., *The Changing Classroom: The Role of the Biological Sciences Curriculum Study*, Doubleday and Company, New York, 1969.
13. Grobman, Hulda, *Evaluation Activities of Curriculum Projects*, AERA Monograph Series on Curriculum Evaluation, No. 2, Rand McNally, Chicago, Illinois, 1968.
14. Grobman, Hulda, *Developmental Curriculum Projects, Decision Points and Processes*, F. E. Peacock, Itasca, Illinois, 1970.
15. Hurd, Paul DeH., *Biological Education in American Secondary Schools, 1890–1960*, No. 1, BSCS Bulletin Series, Biological Sciences Curriculum Study, Boulder, Colorado, 1962.
16. Krathwohl, David; Bloom, Benjamin S.; Masia, Bertram, B., *Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook II, The Affective Domain*, David McKay, New York, 1964.
17. Lindvall, C.M. ; Cox, Richard C., *The IPI Evaluation Program*, AERA Monograph Series on Curriculum Evaluation, No. 5, Rand McNally, Chicago, Illinois, 1970.
18. Lockard, J. David (ed.), *Seventh Report of the International Clearinghouse on Science and Mathematics Curricular Developments, 1970*, a joint project of the Commission on Science Education, American Association for the Advancement of Science and the Science Teaching Center, University of Maryland, College Park, 1970.
19. Mager, Robert F., *Preparing Objectives for Programmed Instruction*, Fearon Press, Palo Alto, California, 1962.
20. McCoy, Eleanor, untitled manuscript of UICSM history, Curriculum Laboratory, University of Illinois, Urbana, forthcoming.
21. Merrill, Richard J.; Ridgway, David W., *The CHEM Study Story*, W. H. Freeman, San Francisco, California, 1969.
22. Scriven, Michael, "The Methodology of Evaluation," in Tyler, Ralph; Gagne, Robert W.; Scriven, Michael, *Perspectives of Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, No. 1, Rand McNally, Chicago, Illinois, 1967.
23. Stake, Robert E. (coord. ed.), *AERA Monograph Series on Curriculum Evaluation*, No. 1—, Rand McNally, Chicago, Illinois, 1967 —. (At the present time, six volumes have been issued, and others are anticipated.)
24. Stake, Robert E., "The Countenance of Educational Evaluation," *Teachers College Record*, 68:523–540, April 1967.
25. Stake, Robert E., "Toward a Technology for the Evaluation of Educational Programs," in Tyler, Ralph; Gagne, Robert; Scriven, Michael, *Perspectives of Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, No. 1, Rand McNally, Chicago, Illinois, 1967.
26. Simpson, Elizabeth, "The Classification of Educational Objectives, Psychomotor Domain," *Illinois Teacher of Home Economics*, Winter 1966–67. Also in Grobman, Hulda, *Developmental Curriculum Projects, Decision Points and Processes*, F. E. Peacock, Itasca, Illinois, 1970.
27. Stufflebeam, Daniel I. and others, *Educational Evaluation and Decision Making*, report of the Phi Delta Kappa National Study Committee on Evaluation, F. E. Peacock, Itasca, Illinois, 1971.
28. Tyler, Ralph, *Basic Principles of Curriculum and Instruction*, University of Chicago Press, Chicago, Illinois, 1950.
29. Tyler, Ralph W. (ed.), *Educational Evaluation: New Roles, New Means*, The Sixty-eighth Yearbook of the National Society for the Study of Education, Part II, University of Chicago Press, Chicago, Illinois, 1969.
30. Wooton, William, *SMSG: The Making of a Curriculum*, Yale University Press, New Haven, Connecticut, 1965.

Research Briefs

THE STATUS OF HUMANITIES PROGRAMS IN OHIO

A GENERAL course in the humanities seems to have been developing for some time. This is a course which crosses subject lines, in either the elementary or secondary school, and contains elements of literature, the social disciplines, and the fine arts. To find out the extent to which it exists, a questionnaire was sent to all school districts in Ohio (634 districts) in fall 1970. It was returned by approximately 60 percent (355) of the districts, and of these almost 16 percent (56) indicated that they had programs in humanities.

An attempt was made to find the extent to which these programs are established, by comparing the number of schools with such programs, with the total number of schools in their districts. Elementary programs in the humanities existed in only about 5 percent of the elementary schools in districts with such programs (these were city school districts). Secondary programs were found in some 41 percent of the high schools in their districts. This degree of establishment was inversely related to district size, with humanities programs being found in smaller percentages of schools as districts get bigger.

Comparison was also made of the degree of financial effort made by districts with programs in the humanities compared to all school districts of the same legal type in the state. Total per pupil expenditures by city school districts with humanities programs exceeded the state average by $28.00. In exempted village and local school

(Continued on page 480.)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Recency of Secondary School Mathematical Preparation as an Influence on First-Year College Mathematics Achievement

ALTON R. GODDARD

ABSTRACT

A random sample of four hundred college freshman mathematics students with comparable high school mathematical preparation was selected from 1,435 students. The test group included 126 members who had not attempted a mathematics course during their last year in high school, designated as Group N, and 263 members who had received mathematical credit during their last year, designated as Group Y. Statistical treatment of data for each group indicated that the SAT mathematical aptitude scores, CEEB (Level I) mathematical achievement scores, and achievement in first-year college mathematics courses was greater for Group Y at the .01 level of confidence. The achievement comparison was based on an analysis of covariance procedure with SAT Mathematical scores, CEEB Mathematical scores, and sex of student used as adjusting variables.

THE BASIC purpose of this study was to determine the effect that recency of mathematical preparation in secondary school has upon achievement in first-year college mathematics courses. This study is primarily concerned with college students who accumulated minimum credit in mathematics in the secondary school. Since a very large portion of this group did not take any type of mathematics course during the last year of secondary school, they are unquestionably affected more by the retention problem than the science-oriented students who normally take mathematics courses during each year of their secondary school preparation. Not all college students major in science or engineering, however, and a very large percent of those who do not are nevertheless required to complete 3 or more semester hours of college credit in the field of mathematics.

Many of the newer multiple-track type curricula in the secondary school have provided high school students a wider range of electives in regard to type and amount of secondary school mathematical preparation. The prospective college student who does not plan to major in science or engineering is normally required to complete only a minimal amount of high school mathematical preparation. The student taking the minimal amount has considerable leeway as to when mathematical courses are taken during secondary school.

At the present time, many of the typical first-year college mathematics classes are composed of a mixture of 8 students varying from those with 4 to 8 semesters of mathematics in their secondary school preparation, to some who had a mathematics course during their last year in secondary school and some who did not. It is not uncommon for a class of this mixture to produce an excessive number of failures.

It is generally assumed that the student with the minimum secondary school mathematical preparation is more likely to be one of the failures. If the group of entering college students with minimum mathematical preparation is examined closely, it seems apparent that the recency of secondary school mathematical preparation is a potentially significant factor in success or failure in a freshman level mathematics course.

Related Research

Several recent studies have been made of the factors affecting achievement in first-year college mathematics courses. The concern in most of these studies was whether the type of mathematical preparation in secondary school was traditional or modern. Wick (9), Coon (2), and Blanton (1) have completed studies indicating there is no significant difference in the quality of preparation of students for first-year college mathematics whether they had the experimental (School Mathematics Study Group Material) or the traditional programs in secondary school mathematics. There seems to be considerable agreement in these studies that the best single predictor of achievement in first-year college mathematics is the secondary school grade record. The particular type of secondary school mathematical preparation of each S is not the primary consideration in the present study but rather the recency of that preparation.

A pioneer study of memory was made by Ebbinghaus (4) in 1913. This research produced a flood of related studies. The experiments of Ebbinghaus have shown that retention is roughly proportional to the degree of learning. This finding has been reported repeatedly in many subsequent experiments.

Levine (7) reports that variation in scores on the mathematical section of the Scholastic Aptitude Test (SAT) reflects the amount of mathematics taken during the senior year. The findings of this research report certainly suggest that a similar relationship could possibly exist between achievement in mathematics courses during the first year in college and the amount of mathematics taken during the senior year in high school.

Douglass (3) points out that it is quite difficult to generalize specifically from studies related to the proportions of high school mathematics once learned which will be remembered after a passage of various amounts of time. He states, however, that it is clear that a large proportion of the forgetting takes place within a very few months. It is to be noted that the critical time unit mentioned is months, rather than years.

The articulation problem in the area of mathematics between high school and college was considered in a study by Kirby (6). The main criticism of the lower-level college mathematics courses is that they are generally traditional and the textbooks being used differ little in content or approach from those of 30 years ago. The findings of his study clearly point out the lack of coordination in the area of mathematics between high school and higher institutions of learning. It is also evident that high school curriculum supervisors are not in a good position to bring about the needed reform in the mathematical philosophies at the college level to produce harmony and compatibility in the area of mathematics. It is noted, however, that if better scheduling can eliminate some of the articulation problem, then the high school personnel have a course of action to pursue until the mathematics courses are more compatible.

Population and Methodology

Subjects in this experiment were students enrolled at Texas A&M University, the oldest public institution of higher education in the State of Texas. The origin of the University is dated by the Morrill Act, which was approved by Congress July 2, 1862. The student enrollment in 1968 was 12,867, including 10,206 undergraduates. The student body is composed predominantly of male students. Female students are admitted under special provisions and at the time of this study comprised 7 percent of the total population.

A total of sixteen acceptable units or credits is required for admission, and the applicant must have graduated from a properly accredited secondary school. Three credits must be in mathematics, two in algebra and one in plane geometry. One-half credit in trigonometry is also required except in liberal arts and teacher education programs, in which cases a substitution of units in foreign language, science, social science, speech, or a vocational subject is permitted. In addition to these requirements, four units of English, two units of science, and two and one-half units of social science are also required.

All Ss in the present study were admitted in the manner previously described. The University permits other types of admissions but none were used in this study.

A total of 1,658 questionnaires were distributed by the assigned instructors to forty-eight classes of college algebra and seventeen classes of plane trigonometry. A total of 1,435 questionnaires were voluntarily completed and returned listing all requested information.

In order to establish a common graduation date, i.e., graduating from high school within 15 days of the same date, all questionnaires were screened and Ss whose graduation date was not 15 days before or after May 25, 1968, were excluded. This action resulted in the rejection of 322 Ss. The remaining 1,113 questionnaires were sorted into two categories. One category included Ss with approximately 3 years of high school mathematical preparation. In general, these Ss had received the same mathematical preparation, including 2 years of algebra and 1 year of plane geometry. No S in this group had completed or attempted a course in analytical geometry, solid geometry, calculus, or any course more sophisticated than perhaps a 1-semester course in elementary analysis.

The remainder of the Ss were rejected and considered unsuitable for being possible candidates of the study group because of their mathematical preparation in secondary school. Most of the rejected Ss had completed a course more sophisticated than plane trigonometry or else they had completed only one semester of mathematics during their senior year in high school. A few Ss were rejected because they had attended summer school at some educational institution prior to the fall of 1968.

The study sample, composed of four hundred Ss, was randomly selected from the group of Ss who had received and completed approximately the same mathematical preparation in secondary school. The random sample was selected utilizing random sampling techniques and procedures described by Snedecor and Cochran (8).

The official grades received by the Ss in the first-year college mathematics courses were collected from the mathematics departmental records. The departmental grade records used were prepared by personnel in the registrar's office at Texas A&M and reflect the transcript records of the Ss composing the test group.

The test group of four hundred Ss included eleven members who had withdrawn from classes, 126 members who had not attempted nor received credit for a mathematics course during their last year in high school—Group N, and 263 members with comparable mathematical preparation who had attempted and received credit for a mathematics course during their last year in secondary school—Group Y.

The basic inferential tool for determining the existence of the differences between the test groups is analysis of variance and analysis of covariance. A computer program written in FORTRAN symbolic language and developed by Gates (5) at the Institute of Statistics, Texas A&M University, was used to calculate the sum of squares and F-values for the analysis of variance and analysis of covariance procedures with unequal subclasses.

Presentation of Data

Information pertaining to the variables (1) mathematics score on the SAT at college entrance, (2) mathematical achievement scores (CEEB Level I) at college entrance, (3) verbal score on the SAT at college entrance, (4) sex, (5) age, and (6) grade received in first-year college mathematics course was collected for each member in the test group. A one-way analysis of variance test was performed on each pair of mean values associated with Group Y and Group N for all variables except sex of student.

The mean values associated with each test group in regard to variables 1, 2, and 6 are shown in Tables 1, 2, and 3. A one-way analysis of variance test was performed on the mean values shown for each test group in Tables 1, 2, and 3. The results of these tests are shown in Tables 4, 5, and 6.

An analysis of covariance test was performed on the mean values of the grade point ratios, shown in Table 3, for Groups Y and N. The results, shown in Table 7, were determined while treating the variables—(1) mathematical aptitude score on the SAT at college entrance; (2) mathematical achievement score (CEEB Level I) at college entrance; and (3) sex of student—as adjusting variables.

Table 1.—Mean Mathematical Aptitude Test Scores of the SAT for Students Classified According to Recency of High School Mathematical Preparation

	Group Y	Group N
Mathematical Aptitude Scores	546.72	504.86
N	263	126

Table 2.—Mean Mathematical Achievement Scores of the CEEB Mathematics Achievement Test (Level I) for Students Classified According to Recency of High School Mathematical Preparation

	Group Y	Group N
CEEB Level I Achievement Scores	534.98	462.18
N	263	126

Table 3.—Mean Grade Point Ratios of Grades Received in First-Year College Mathematics Courses for Students Classified According to Recency of High School Mathematical Preparation

	Group Y	Group N
Grade Point Ratio	2.94	1.48
N	263	126

Findings and Conclusions

Mathematical Aptitude Score of the SAT. The mean value of the mathematical aptitude test scores of the SAT for Group Y was statistically greater than the similar mean value for Group N. The difference between the means was significant at the .001 level of confidence, as shown in Table 4.

Verbal Scores of the SAT. The difference be-

Table 4.—Analysis of Variance for Mathematical Aptitude Scores of the SAT

Source of Variation	df	Mean Square	F
Recency of High School Mathematical Preparation	1	149,282.00	26.09*
Within	387	5,720.96	

$*p < .001$

Table 5.—Analysis of Variance for Mathematical Achievement Scores of the CEEB Mathematics Achievement Test (Level I)

Source of Variation	df	Mean Square	F
Recency of High School Mathematical Preparation	1	451,509.00	96.96*
Within	387	4,656.78	

*$p < .001$

tween the mean value of the verbal scores of the SAT for Groups Y and N was not statistically significant at the .05 level of confidence.

Sex of Students. The number of female students in each test group was approximately the same and was inadequate in size to be considered a respectable statistical sample. The findings reported in this study are based essentially on a male population.

Age of Students. The difference between the mean age for both test groups was not statistically significant at the .05 level of confidence. The characteristic age for both test groups was 18 years.

Grade Point Ratios of Grades Received in First-year College Mathematics Courses. The mean value of the grade point ratio for Group Y was statistically greater than the mean value of the grade point ratio for Group N. The difference was statistically significant at the .001 level of confidence, as shown in Table 5.

Mathematical Achievement Scores (CEEB Level I). The mean value of the scores on mathematical achievement (CEEB Level I) for Group Y was statistically greater than the similar mean value for Group N at the .001 level of confidence as indicated in Table 6.

Table 6.—Analysis of Variance for Grades Received in First-Year College Mathematics Courses

Source of Variation	df	Mean Square	F
Recency of High School Mathematical Preparation	1	209.61	183.87*
Within	387	1.14	

*$p < .001$

Grades Received in First-year College Mathematics Courses With Adjustment Made for all Covariates Determined Statistically Significant in this Study. The mean value of the grade point ratios for grades received in first-year college mathematics courses by members of Group Y was statistically greater than the similar mean value for members of Group N at the .001 level of confidence as shown in Table 7. This difference was found after making adjustments for existing differences between the test groups in regard to variation in (1) mathematical aptitude scores on the SAT, (2) mathematical achievement scores (CEEB Level I), and (3) sex of student.

Recommendations

The prospective college student should schedule mathematical preparation in a manner which would prevent a time lapse of one or more semesters between completion of his final course in high school and his first college course. High school students desiring to make optimal use of their mathematical preparation in high school should, if required for college entrance, take the

Table 7.—Analysis of Covariance for Grades Received in First-Year College Mathematics Courses

Source of Variation	df	Mean Square	F
Independent Variable			
Recency of High School Math Preparation	1	114.21	109.50**
Covariates			
Scholastic Aptitude Test			
Mathematical Aptitude Scores	1	3.21	3.07
CEEB Level I Mathematical Achievement Scores	1	13.07	12.54*
Sex of Students	1	7.23	6.93
Within	384	1.04	

**$p < .001$

*$p < .01$

College Entrance Board's SAT and Mathematics Achievement Test (Level I) as soon as possible after the completion of the major portion of the last course in high school mathematics. The entering college student with the minimum high school mathematical preparation should commence required college level mathematical training during the first semester enrolled in college.

REFERENCES

1. Blanton, Floyd Lamar, "The Effect of Having Followed Certain Programs in High School Mathematics on Achievement in the First Course in College Mathematics," unpublished doctoral dissertation, The University of Georgia, Athens, 1963.
2. Coon, Lewis Hulbert, "School Mathematics Study Group Mathematics as a Factor Influencing Success in Freshman Calculus," unpublished doctoral dissertation, Oklahoma State University, Stillwater, 1963.
3. Douglass, H. R., "Permanency of Retention of Learning Secondary School Mathematics," *The Mathematics Teacher*, 29:287–288, 1936.
4. Ebbinghaus, H., *Memory*, Teachers College, Columbia University, New York, 1913.
5. Gates, C. E., "General Linear Hypothesis," unpublished paper, Institute of Statistics, Texas A&M University, College Station, 1968.
6. Kirby, William Abbott, "A Critical Analysis of Certain Collegiate Mathematics Programs in Texas, Oklahoma, and Arkansas," unpublished doctoral dissertation, The University of Texas, Austin, 1963.
7. Levine, Richard S.; Angoff, William H., *The Effects of Practice and Growth on Scores on the Scholastic Aptitude Test*, Educational Testing Service, Princeton, New Jersey, 1958.
8. Snedecor, George W.; Cochran, William G., *Statistical Methods*, The Iowa State University Press, Ames, 1967.
9. Wick, Marshall Eldon, "A Study of Factors Associated with Achievement in First-Year College Mathematics," unpublished doctoral dissertation, The University of Minnesota, Minneapolis, 1963.

New Books from the Publishers

Children

Edge, David. *The Formative Years: How Children Become Members of Their Society* (New York, NY: Schocken Books, 1970), xvi + 72 pp. Hardcover $4.50. Paper $1.75.

Kujoth, Jean S. *Reading Interests of Children and Young Adults* (Metuche, NJ: Scarecrow Press, Inc., 1970), xiii + 449 pp. Hardcover $10.00.

Middleton, John (ed.). *From Child to Adult—Studies in Anthropology of Education* (Garden City, NY: The Natural History Press, 1970), xx + 355 pp. Hardcover $8.95. Paper $3.95.

Education

Billett, Roy O. *Improving the Secondary-School Curriculum: A Guide to Effective Curriculum Planning* (New York, NY: Atherton Press, 1970), xxvii + 364 pp. Hardcover $8.50.

Cowan, L. Gray. *The Cost of Learning: The Politics of Primary Education in Kenya* (New York, NY: Teachers College Press, 1970), xiii + 106 pp. Paper $4.50.

Dussault, Gilles. *A Theory of Supervision in Teacher Education* (New York, NY: Teachers College Press, 1970), xv + 275 pp. Paper $3.50.

Epstein, Herman. *A Strategy for Education* (New York, NY: Oxford University Press, 1970), vii + 122 pp. Hardcover $4.95.

Foshay, Arthur (ed.). *The Professional as Educator* (New York, NY: Teachers College Press, 1970), viii + 128 pp. Paper $4.95.

Li, Y. T. *Education in Creative Engineering* (Cambridge, MA: The MIT Press, 1970), xxxiii + 386 pp. Paper $6.95.

Niblett, W. R. *All for the Best?* (Great Britain: Geo. G. Harrap and Co. Ltd., 1970), 27 pp. Paper 30 pence.

Riesman, D. Gusfield and J. Gamson. *Academic Values and Mass Education* (Garden City, NY: Doubleday and Co., Inc., 1970), xxii + 332 pp. Hardcover $7.95.

The Scottish Council for Research in Education. *Forty First Annual Report 1968–69* (46 Moray Place, Edinburgh 3), 39 pp. Free.

Unesco. *International Yearbook of Education* Vol. XXXI–1969 (Geneva: United Nations Educational, Scientific and Cultural Organization, 1970), xxiii + 235 pp. Paper $6.50.

Research

Ashlock, Robert and Wayne Herman. *Current Research in Elementary School Mathematics* (New York, NY: The Macmillan Co., 1970), xiii + 497 pp. Paper $5.95.

Ballou, Stephen. *A Model for Theses and Research Papers* (Boston, MA: Houghton Mifflin Co., 1970), vii + 70 pp. Paper $1.00.

Newmann, Fred M. and Donald W. Oliver. *Clarifying Public Controversy: An Approach to Social Studies* (Boston, MA: Little, Brown and Company, 1970), vii + 356 pp. Hardcover $6.95.

Nisbet, J. D. and N. J. Entwistle. *Educational Research Methods* (New York, NY: American Elsevier Publishing Co., Inc., 1970), 192 pp. Hardcover $7.25.

Sukhatme, Pandurang V. and Balkrishna V. Sukhatme. *Sampling Theory of Surveys with Applications* (Ames, IA: Iowa State University Press, 1970), xvi + 452 pp. Hardcover.

Teaching

Corcoran, G. *Language Arts in the Elementary School* (New York, NY: The Ronald Press Company, 1970), v + 280 pp. Hardcover $5.50.

Haddan, Eugene. *Evolving Instruction* (New York, NY: The Macmillan Company, 1970), xiii + 335 pp. Hardcover $6.95.

Hook, J. N. and Michael G. Crowell. *Modern English Grammar for Teachers* (New York, NY: The Ronald Press Company, 1970), v + 291 pp. Hardcover $5.50.

MacKenzie, Norman; Michael Eraut and Hywel C. Jones. *Teaching and Learning: An Introduction to New Methods and Resources in Higher Education* (UNESCO: and The International Assoc. of Universities, 1970), 209 pp. Paper $3.50.

May, Lola J. *Teaching Mathematics in the Elementary School* (New York, NY: The Free Press, 1970), xiii + 317 pp. Hardcover $7.95. Paper $4.95.

Ryan, Kevin. *Don't Smile Until Christmas* (Chicago, IL: The University of Chicago Press, 1970), xiii + 190 pp. Hardcover $5.95.

Slatoff, Walter J. *With Respect to Readers* (Ithaca, NY: Cornell University Press, 1970), viii + 211 pp. Paperback $5.95.

Smith, William I. *Guidelines to Classroom Behavior* (Brooklyn: Book-Lab, Inc., 1970), 128 pp. Paper $2.50.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

The Interrelationships Among Self-Esteem, Personal Values, and Interpersonal Values[1]

SANDRA LOU STEIN, University of South Carolina

ABSTRACT

The interrelationships of self-esteem, personal values, and interpersonal values among high school students were considered. Coopersmith's Self-Esteem Inventory and Gordon's Survey of Personal Values and Survey of Interpersonal Values were administered to 1,720 high school students in northern Illinois. Data were analyzed by Pearson product-moment correlation coefficients. No significant relationships existed between self-esteem and each of the six personal values and between self-esteem and each of the six interpersonal values with one exception. Self-esteem correlated significantly with the value of Achievement for seniors. The lack of expected relationships may be due to deficiencies in the instruments used for measurement or to the individuality of value and self-esteem combinations.

THE PURPOSE of the study was to investigate relationships between self-esteem and personal values and self-esteem and interpersonal values in high school students. In the past years, there has been increasing evidence that learning depends on not only good teaching, but also the students' values and self-evaluations. Knowledge of the relationship of these variables and normative data on them may provide educators with suggestions for curricular or classroom changes. It may provide information on the current status and relationship of these variables for those wishing to instill or make changes in the self-esteem or values of our youth. Further, it may add to the general fund of knowledge on adolescent development.

Although numerous studies have been made on various aspects of the self, there has been no agreement on definition of terms or techniques of measuring it. Wylie concluded that studies of self-esteem have been more fruitful than studies of broader terms such as self-concept (6). The present study focuses on self-esteem as the evaluative aspect of the self-concept. The Self-Esteem Inventory (SEI) developed by Coopersmith, was chosen as the instrument to measure self-esteem, and his definition of self-esteem as being "a *personal* judgment of worthiness that is expressed in the attitudes the individual holds toward himself" is used as the basic frame of reference (1). The instrument has previously been used with a high school population and has received a very favorable review from Robert R. Sears (5).

There has also been a paucity of standardized instruments to measure values at the high school level. A pioneer instrument in the measurement of values was the Allport–Vernon Study of Values which was designed for college students. Recently, Leonard V. Gordon developed two instruments which should facilitate the study of values at the high school level. His Survey of Interpersonal values (SIV) was designed to measure "values involving the individual's relationships to other people or their relationships to him" (2). He subsequently developed the Survey of Personal Values (SPV) to measure "values that help determine the manner in which an individual copes with the problems of everyday living" (3). These instruments appeared to be appropriate measures and were used in the study.

Design of the Study

The study was an *ex post facto* study, as defined by Kerlinger (4). The hypothesis was stated in null form to facilitate statistical testing. The .05 level was used as a criterion for significance. The hypothesis was:

Self-esteem does not significantly correlate with any of the personal values or with any of the interpersonal values: (a) Self-esteem does not significantly correlate with any of the personal or interpersonal values by sex; (b) Self-esteem does not significantly correlate with any of the personal or interpersonal values by grade.

Sample. A total of 1,720 students from two high schools in northern Illinois were tested. Of these students, 128 invalidated their tests and were omitted from the sample. The remaining 1,592 questionnaires were used to study the relationships of self-esteem to personal and interpersonal values.

Procedures. The SEI (1), the SPV (3), and the SIV (2) were administered to all students midway through the academic year. Directions for all three scales were read to the students by the writer and/or one of eight graduate students and one mature adult. The students were then given a 45-minute period in which to complete all three questionnaires.

On the SEI, the students were asked to mark fifty statements as being like or unlike them. The directions on the SPV and SIV were to mark the item most important to them and least important to them in each triad. Background information was collected on a separate data sheet.

Instruments

Self-Esteem Inventory. The SEI, a 50-item self-report instrument which yields a total score that may range from 0 to 100 points, was developed and originally used by Coopersmith (1). A high score indicates high self-esteem. Coopersmith reports a test-retest reliability of .88 after 5 weeks with a sample of thirty children and of .70 after 3 years with a sample of fifty-six children.

Survey of Personal Values. The SPV, developed by Gordon (3), consists of thirty triads within which the student must select the one most important to him and the one least important to him. A high score indicates that the student is strongly oriented toward that value.

Test-retest reliabilities were determined for a sample of fifty-eight college males and thirty-nine college females with 7 to 10 days between administrations (.84). Kuder-Richardson reliabilities were determined based on the responses of 167 college students (.83).

Survey of Interpersonal Values. The SIV, also developed by Gordon (2), also consists of thirty triads within which the student must select the one most important to him and the one least important to him. A high score indicates that the student is strongly oriented toward the value.

Test-retest reliabilities were determined for a sample of seventy-nine college students with 10 days between administrations (.88). Kuder–Richardson reliabilities were established from data on a sample of 186 college students (.83).

Results and Discussion

Pearson product-moment correlation coefficients were calculated for the population as a whole, for each sex, and for each grade level in order to study the relationship of self-esteem to each of the six personal values and to each of the six interpersonal values. These correlations are presented in Table 1. For seniors, self-esteem correlated to the degree of .21 ($p < .05$) with the personal value of Achievement. All other correlation coefficients for the personal values and the interpersonal values were not observed to be significantly different from zero.

The following conclusion can be drawn regarding the hypothesis of this study: ***With one exception, there were no statistically significant correlations between self-esteem and values.*** There was a significant, positive relationship between self-esteem and the personal value of Achievement for seniors. It seems logical that by the last year in high school, an increase in self-esteem would be associated with an increase in the students' value of achievement. By the senior year

Table 1.—Correlations of Self Esteem with Personal Values and with Interpersonal Values (by Combined Group, Sex, and Grade)

	Personal Values						Interpersonal Values					
Self-Esteem	P	A	V	D	O	G	S	C	R	I	B	L
Combined	-.075	.145	-.156	.032	.003	.163	-.074	.075	-.047	-.093	.058	.113
Sex: M	-.092	.150	-.154	.041	-.003	.144	-.104	.085	-.106	-.086	.107	.079
F	-.059	.140	-.156	.026	.011	.181	-.038	.066	.013	-.111	.027	.153
Grade: 9	-.012	.155	-.194	.082	.010	.188	.003	.061	-.028	.013	.021	.055
10	-.070	.077	-.103	-.011	-.034	.149	-.075	.140	-.050	-.178	.124	.059
11	-.065	.176	-.163	-.043	-.012	.166	-.124	.105	-.071	-.147	.058	.142
12	-.170	.208*	-.178	.119	-.026	.141	-.140	.076	-.064	-.161	.073	.191

* Significant at .05 level. (.r of ± .195 needed for significance at .05 level.)

many students who may not value achievement at all have dropped out of school, and the students who are college-bound have seen the need for achievement for college entrance. Other students may place less value on achievement (which is connected to school) and look for rewards outside the school system.

The study did not demonstrate that self-esteem and values are highly related. This may be a true observation, or it may be an artifact of the instruments measuring self-esteem and values. Perhaps different instruments, which measure these same values or other values, might be more appropriate measures and therefore show some relationship to self-esteem. It may be that due to the conceptual limitations of students at this age, they cannot reflect their actual values on paper-and-pencil tests, and other ways of assessing values of teenagers must be devised.

However, it may be that self-esteem is not associated with any particular set of values but may vary according to one's success with his own peculiar combination of values. Thus, while values could affect self-esteem, or vice versa, this relationship could not be established for groups, since every individual would have his own unique combination of values.

FOOTNOTE

1. This research is adapted from the author's doctoral dissertation at Northern Illinois University. The author gratefully acknowledges her indebtedness to Marvin Powell who served as advisor.

REFERENCES

1. Coopersmith, Stanley, *The Antecedents of Self-Esteem*, W. H. Freeman and Company, San Francisco, California, 1967.
2. Gordon, Leonard V., *Survey of Interpersonal Values Manual*, Science Research Associates, Inc., Chicago, Illinois, 1960.
3. Gordon, Leonard V., *Survey of Personal Values Manual*, Science Research Associates, Inc., Chicago, Illinois, 1967.
4. Kerlinger, Fred N., *Foundations of Behavioral Research*, Holt, Rinehart and Winston, Inc., New York, 1964.
5. Sears, Robert R., "A Treatment of Worth," *Contemporary Psychology*, 14:146–147, March 1969.
6. Wylie, R. C., *The Self Concept*, University of Nebraska Press, Lincoln, 1961.

Forthcoming Articles

Herr, Edwin L.; Pennsylvania State University, University Park
"STUDENT NEEDS, COLLEGE EXPECTATIONS, AND 'REALITY' PERCEPTIONS"

Lemke, Elmer A. and James T. Hecht; Illinois State University, Normal
"EFFECTS OF A DEGREE OF TRAINING, GROUP SIZE, AND INDUCTIVE ABILITY ON THE TRANSFER OF CONCEPTUAL BEHAVIOR"

Menges, Robert J. and Robert Marx; University of Illinois, Urbana
"ACHIEVEMENT OF HIGH-RISK AND REGULAR STUDENTS IN TRADITIONAL AND INNOVATIVE PSYCHOLOGY COURSES"

Mickelson, Norma I.; University of Victoria, Victoria, B. C., Canada
"MEANINGFULNESS INDICES FOR 120 MASS AND COUNT NOUNS FOR CHILDREN AGED 14 YEARS AND FOR UNIVERSITY STUDENTS"

Morrison, Coleman; Temple University, Philadelphia, Albert J. Harris; City University of New York and Irma T. Auerbach; Harvard University
"THE READING PERFORMANCE OF DISADVANTAGED EARLY AND NON-EARLY READERS FROM GRADES 1 THROUGH 3"

Quilling, Mary and Wayne Otto; Wisconsin Research and Development Center for Cognitive Learning, Madison
EVALUATION OF AN OBJECTIVE BASED CURRICULUM IN READING

Shapiro, Bernard J.; Boston University, Robert E. Willford; Radiant Corporation, and Phyllis P. Shapiro; Emmanuel College, Boston
"THE i.t.a. TRANSITION PROBLEM—A COMPARATIVE STUDY"

Shaver, James P.; Utah State University, Logan and Hyrum E. Richards; Casper-Midwest Schools, Casper, Wyoming
"OPEN-CLOSED MINDEDNESS AND AN INQUIRY-ORIENTED SOCIAL STUDIES METHODS COURSE"

Shepps, R. Ronald and Florence P.; Kew Gardens Hills, New York
"RELATIONSHIP OF STUDY HABITS AND SCHOOL ATTITUDES TO ACHIEVEMENT IN MATHEMATICS AND READING"

Yens, David P.; Pennsylvania State University, University Park
"EVALUATION OF AN INNOVATIVE EDUCATIONAL DEVICE—A CASE STUDY"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

A Longitudinal Study of Visual Perceptual Training and Reading Achievement[1]

NEAL R. GAMSKY and FAYE WILLIAMS LLOYD

ABSTRACT

The impact of the Frostig Program for the Development of Visual Perception on students participating in the study was examined. The study involved ten teachers and twenty kindergarten classes from six rural school districts. Pre- posttests with the Frostig and Metropolitan Readiness Tests were administered to control and experimental groups. The experimental group received the Frostig program for 15 minutes each day for 4½ months. All groups were tested on the Stanford Achievement Test the following year. Results indicated that the Frostig program benefited the students' reading ability since achievement test scores of those using the program exceeded scores of those not participating in it.

A REVIEW of the literature indicates that for the past decade there has been great concern regarding the relationship of specific visual perception capabilities to reading achievement. Evidence has been presented with respect to the effectiveness of the Frostig Developmental Test of Visual Perception (DTVP) (3) in predicting reading readiness in kindergarten (1) as well as to significant relationships between this instrument and reading achievement test scores in later grades (6, 7). However, while some have discovered that children using the Frostig Visual Perception Training Program (VPTP) did better on the DTVP than controls (4), it has also been found that the training did not produce actual gains in reading (2, 5, 8). In view of the importance of this subject especially for the early childhood years, further research with the VPTP would appear to be warranted.

Purpose and Procedure

The purpose of this study was to answer the following questions:

1. *Do children in kindergarten who take the Frostig program achieve better than controls on the Frostig Visual Perception Test?*
2. *Do kindergarten children who used the Frostig program achieve higher on the Metropolitan Reading Readiness Test (MRT) at the end of their kindergarten year as compared to controls?*
3. *Does the DTVP predict which children will later have difficulty in reading?*
4. *Do children using the VPTP achieve higher on reading tests as compared to controls?*

Subjects used in this study were from six Midwestern rural school districts. They attended ten different schools and were in intact classes. Selection was determined by the availability of teachers and administrators who agreed to participate. The twenty kindergarten classes were divided into two groups and were designated as experimental and control groups. Both groups were randomly selected from morning and afternoon sessions of the same teacher thereby controlling teacher and time variables. All groups received the Frostig DTVP as a pretest and posttest, and the MRT was administered at the end of the kindergarten year. All children in the experimental group received the VPTP for 15 minutes each day for 4½ months from midyear to May. The following year (1968–69), pupils from both groups were assigned as usual to classes throughout the school districts. At the end of the academic year all students were given the Stanford Achievement Test Primary I Battery Form W.

Mean group differences between Frostig DTVP gain scores, MRT scores and Stanford Achievement scores were analyzed by analysis of variance and analysis of covariance techniques. Tukey's q-test was used to determine differences between means. Main effects were treatment, sex, and perceptual quotient (PQ) quartile blocks (Q_1 = PQ < 90, Q_2 = PQ 90–99, Q_3 = PQ 100–109, Q_4 = PQ 110–125). The F-test and scatter grams were used to test the assumptions of

Table 1.—Analysis of Variance Summary: Frostig Test of Visual Perception Treatment Groups as Main Effect of Factorial Analysis

	Experimental Gain Score $\bar{X}$ N=10 Classes	Control Gain Score $\bar{X}$ N=10 Classes	Total Variance	Error Between	df	Error Within	df	F-Ratio	Prob.
(1) Eye Motor Coordination	1.1348	.7067	3.025	5.856	1	3.089	114	1.8959	.167
(2) Figure Ground	2.1739	1.2603	3.757	26.503	1	3.521	114	7.5274	.007**
(3) Constancy of Shape	2.7208	1.8735	7.357	23.678	1	5.889	114	9.3571	.040*
(4) Position in Space	2.2783	1.1376	4.850	41.508	1	4.418	114	9.3946	.003**
(5) Spatial Relationships	1.7333	1.0050	3.019	16.653	1	2.716	114	6.1308	.010**
(6) Perceptual Quotient	15.0811	4.0606	127.012	4101.325	1	80.409	114	51.0057	.000**

* $p < .05$
** $p < .01$

homogeneity of variance and linearity respectively. Both assumptions were satisfied. All analyses were computed via Univac 1108[2] with a calculator complement.

Results

Question 1: Did children in kindergarten who took the Frostig program achieve better than controls on the Frostig Visual Perception Test?

A summary of the Frostig test results is presented in Tables 1 and 2. Frostig VPTP class means showed significant gains over control class means on subtests 2–figure-ground perception, 4–perception of position in space, 5–perception of spatial relationships and the PQ ($p < .01$), and on subtest 3–perceptual constancy ($p < .05$); but there was no significant gain on subtest 1–visual-motor coordination. Significant differences between sexes were evident on subtests 3, 5, and the PQ. Quartiles 2 and 4 made significantly greater gains than quartiles 1 and 3 on perceptual constancy and the PQ (see Table 2).

Except for subtest 4–perception of position in space ($p < .12$), almost none of the variation between group gains can be attributed to practice effect of pretest or prior knowledge ($p < .05$).

Question 2: Did children who used the Frostig program achieve higher scores on the MRT at the end of their kindergarten year as compared to controls?

Data with regard to this question are presented in Table 3. Pupils receiving the Frostig VPTP scored significantly higher on all Metropolitan Achievement scales except listening ($p < .01$). Also, girls scored significantly higher than boys on all scales except word meaning.

Metropolitan scales 2 and 3 (listening and matching) did not attribute variation between groups to the Frostig DTVP treatment; however,

Table 2.—Results on the Frostig Test of Visual Perception and Composite Gain Score Comparisons for Experimental and Control Groups

Subtest	Treatment	Sex	Quartiles△	T x S
(1) Visual-motor coordination				
(2) Figure-ground perception	**	.06		
(3) Perceptual constancy	**	**	.0001	.03
(4) Perception of position in space	**			
(5) Perception of spatial relationships	*	**		
(6) PQ	**	**	**	

*$P < .05$
**$P < .01$

△$[Q_4 > Q_2] > Q_1$ and Q_3 ($P < .05$)

Quartile determined by PQ score	
1	0-89
2	90-99
3	100-109
4	110-125

scales 1, 4, 5, and 6 (word meaning, alphabet, numbers, and copying) attributed variation between groups to level of attainment on the Frostig posttest.

Question 3: Did the DTVP predict which children later have difficulty in reading?

Question 4: Did children using the VPTP achieve higher on reading tests as compared to controls?

The last two research questions will be dealt with together. A summary of the results of the analysis is presented in Table 4. Pupils receiving the Frostig VPTP scored significantly higher on Stanford Achievement scales—1 word reading, 2 paragraph meaning, 4 spelling, 5 word study skills, and 6 arithmetic ($p < .01$)—but not on 3 vocabulary ($p < .10$). Except for scale 5—word study skills ($p < .06$), significant differences between sexes were evident ($p < .05$). In each group girls always scored higher than boys. There was no significant difference between posttest quartiles on Stanford scores.

The variation between treatment groups was attributable on all Stanford Achievement scales except arithmetic to level of attainment on Frostig posttest scores. Experimental PQ < 90 children scored consistently higher than non-treatment children on the Stanford.

The Frostig Test of Visual Perception appeared to be useful in predicting which children will have difficulty in reading since non-treatment PQ < 90 youngsters ranked consistenty low on the majority of subtests while experimental youngsters ranked consistently higher on all subtests except arithmetic. The significant correlation of Frostig posttest scores with treatment

Table 3.—Summary of Factorial Analysis of Variance and Covariance Results on Ss' Metropolitan Achievement Test Scores

	A Treatment $\bar{X}$'s			B Sex $\bar{X}$'s			C Post-Test Quartile	AB	AC	BC	ABC	Covariance Adjusted with Frostig Post-Test Scores
Metropolitan Achievement Scales	X N=10 Classes	C N=10 Classes		M	F							
(1) Word Meaning	11.3503	10.1553	**	10.4374	11.0682	NS						NS
(2) Listening	11.4437	10.9261	.20	10.7330	11.6368	*				**	*	*
(3) Matching	10.6580	8.0341	**	8.1763	10.5158	**						*
(4) Alphabet	10.7813	9.0305	**	9.0985	10.7133	**						NS
(5) Numbers	15.5692	12.8301	**	13.4055	14.9937	*						NS
(6) Copying	9.2246	6.6955	**	7.0222	8.8979	**						NS

* $p<.05$
** $p<.01$
NS = Not Significant

Table 4.—Summary of Factorial Analysis of Variance and Covariance Results on Ss' First Grade Stanford Achievement Test Scores

	A Treatment $\bar{X}$'s			B Sex $\bar{X}$'s			C Post-Test Quartile	AB	AC	BC	ABC	Covariance Adjusted with Frostig Post-Test Scores
Stanford Achievement Scales	X N=10 Classes	C N=10 Classes		M	F							
(1) Word Reading	20.0270	16.5838	**	17.1292	19.4815	**						NS
(2) Paragraph Meaning	19.8299	16.6398	**	17.0959	19.3739	**						NS
(3) Vocabulary	23.7114	21.2720	.10	20.1671	24.8163	**						NS
(4) Spelling	21.0787	17.1259	**	17.3836	20.8220	**						NS
(5) Word Study Skills	24.9159	18.8049	**	20.1289	23.5919	.06						NS
(6) Arithmetic	19.9328	17.0374	**	16.9585	20.0117	**						*

* $p<.05$
** $p<.01$
NS = Not Significant

groups' scores to account for the difference between groups would seem to be an indication that level of attainment on the DTVP may be a means to predict reading achievement scores. It likewise appeared that the greater achievement scores of the treatment group were due to the effects of the training program.

Discussion and Conclusions

In general, the findings of this study lend support to other studies (4, 8) indicating that the Frostig VPTP benefits children by improving visual perceptual abilities. Both treatment groups made significant gains in visual perceptual abilities ($p < .05$), but the treatment group made greater gains. Control group gains may have been due to pupil maturation or the normal efforts of teachers in the course of routine classroom activities.

It appears that the Frostig Test of Visual Perception is useful in predicting which children will have difficulty in reading development. Youngsters who had a PQ of 90 or less (Q_1) ranked persistently lowest on achievement while those with a PQ of 110–125 (Q_4) were persistently highest achievers on the Metropolitan Achievement Test. Apparently the program does not have its greatest impact on those most in need of visual perceptual gain. PQ level appears to predict degree of gain in visual perceptual abilities. Youngsters with PQ's less than 90 made less PQ gain than youngsters with PQ's over 90. A high correlation between PQ and IQ may mean that we are actually measuring effects of intelligence which are known to be directly related to achievement. If this is the case, follow-up studies comparing perceptual pretest scores and gain scores, and an analysis of covariance comparing adjusted achievement scores with IQ as an outside criterion are in order. It appears that sex is a major factor in visual perceptual growth. Girls made greater gains than boys on all Frostig subtests and the PQ. Since it is known that girls are better achievers than boys in the early school years, this is not a surprising outcome.

In contrast to other studies (5, 8), the findings of this study also support the belief that the Frostig VPTP helps reading ability since achievement test scores of those using the training program exceed those not participating in it.

FOOTNOTES

1. This research was conducted as a part of the evaluation of a Title III (ESEA) project #67–03204–1, "Exemplary Programs of Pupil Services in Small Towns and Rural Areas," in CESA #13, Waupun, Wisconsin. The cooperation of the school administrators in CESA #13 is gratefully acknowledged.
2. Dixon, W. J., ed., "Analysis of Variance for Factorial Design" and "Analysis of Covariance with Multiple Covariates," *BMD BioMedical Computer Programs*, Science Computing Facility, University of California, Los Angeles, pp. 495–542.

REFERENCES

1. Bryan, Quentin, "Relative Importance of Intelligence and Visual Perception in Predicting Reading Achievement," *California Journal of Educational Research*, 15:44–48, January 1964.
2. Cohen, S. Alan, "Studies in Visual Perception and Reading in Disadvantaged Children," *Journal of Learning Disabilities*, 2:498–507, October 1969.
3. Frostig, M.; Lefever, W.; Whittlesey, J., *Developmental Test of Visual Perception* (3rd Ed.), Consulting Psychologist Press, Palo Alto, California, 1964.
4. Jacobs, James N., "An Evaluation of the Frostig Perception Training Program," *Educational Leadership*, 15:332–340, January 1968.
5. Jacobs, James N.; Wirthlin, Lenore D.; Miller, Charles B., "A Follow-Up Evaluation of the Frostig Visual-Perceptual Training Program," *Educational Leadership Research Supplement*, pp. 169–175, November 1968.
6. Olson, Arthur V., "Relation of Achievement Test Scores and Specific Reading Abilities to the Frostig Developmental Test of Visual Perception," *Perceptual and Motor Skills*, 20:179–184, 1966.
7. Olson, Arthur V., "School Achievement, Reading Ability, and Specific Perception Skills in the Third Grade," *The Reading Teacher*, pp. 490–492, April 1966.
8. Rosen, Carl L., "An Experimental Study of Visual Perceptual Training and Reading Achievement in First Grade," *Perceptual and Motor Skills*, 20:979–986, 1966.

BOOK REVIEWS

ROBERT E. CLASEN
Book Review Editor

Private Wealth and Public Education

Coons, John E.; Clune, William H. III; Sugarman, Stephen D. Harvard University Press, Cambridge, Massachusetts, 1970, xxvi + 520 pp. Hardcover $12.50.

PRIVATE WEALTH and Public Education is a most timely publication. In recent years contributions through foundations and individuals which are not from taxes have greatly increased. In the first place, money in the hands of private foundations for distribution has increased. In addition, many school districts attempting to avoid desegregation have established nonpublic schools which are supported by contributions and tuition. Furthermore, there has developed a great inequality of financing in spite of the attempts of states to insure equality.

Part I of this book deals with the examination of the existing financial structure in a number of states and shows how the various state aid formulas fail to eliminate inequities caused by differences in taxable wealth per child in different localities.

A proposed solution to eliminate, or at least greatly reduce, the inequalities is set forth in Part II formulated in eighty-six pages. Part III is an attempt to evaluate the proposition which is a proposal of the authors.

The book is written in a most scholarly style. Footnotes and quotations indicate the very great reading on the part of the authors. The book is written more for the scholars who wish to go into this question than for the general reader.

Harl R. Douglass
Consultant on Education
Boulder, Colorado

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Effects of Same Sex Class Organization on Junior High School Students' Academic Achievement, Self-Discipline, Self-Concept, Sex Role Identification, and Attitude Toward School[1]

JOSEPH R. ELLIS and JOAN L. PETERSON
Northern Illinois University

ABSTRACT

The purpose of the study was to determine the effect of same-sex class organization on junior high school (seventh and eighth grade) students' academic achievement, self-discipline, self-concept, sex role identification, and attitude toward school. A pretest and posttest experimental design with randomly assigned groups (with some adjustments in assignment) was employed. Each experimental group was in English, social studies, mathematics, science, and physical education classes for a full school year. The instructional content and method were essentially the same for both the control and experimental groups. Five null hypotheses, one for each dependent variable, were tested by application of analysis of covariance using a 2×2×2 factorial design. While the data did not warrant the rejection of any of the null hypotheses, significant differences were found for factors not effected by the experimental treatment. Differences in academic achievement and general maturity were consistent with the research about early adolescence in that the differences favored the girls.

IN AMERICA, early adolescence is often fraught with rapid change and much frustration for the individual. For most, it is a time for breaking away from childhood and embarking upon new behavior patterns and roles that call for considerable adaptations. This period is characterized by changes which confront the individual with new developmental tasks in the area of his or her intellectual, social, emotional, moral, self, and physical development. The effectiveness of learning to cope with these tasks influences subsequent behavior and may leave an enduring impression on the personality (3).

One continuing concern of those working with adolescents is the difference in maturity between boys and girls at this age. It is generally assumed that the year to year-and-one-half advance in maturity which girls have over boys at this level works to the detriment of boys, especially when boys must associate and compete with girls in an environment that is generally dominated by feminine values and behavior. Sex differences seem to be most intense during adolescence (4). Does the association and competition between boys and girls in the classroom at the seventh and eighth grade level result in significant differences in the development of boys and girls? It is to this question that this study was addressed.

American educators have expressed a great deal of concern about meeting individual differences in the classroom, but the most significant individual differences of all—sex of the student—has received little attention (4). Instead, coeducational classroom grouping has been taken for granted in American schools, although no empirical evidence exists to show that it is superior to single-sex grouping in meeting any of the educational purposes assumed by American schools. When questioned at all, coeducational classroom organization has been justified by statements that it is "realistic," economically advantageous, a wholesome socializing experience for both sexes, and the best way to provide equal educational opportunity for women (11). Margaret Mead pointed out the problems presented by coeducation:

> *This relative advancement lasts through college, and it isn't until graduate school that the boys catch up with girls. We are dealing with a serious discrepancy here and our nation of coeducation by chronological age is not coming to terms with the biological facts (6).*

Grambs and Waetjen have pointed out that while educators have professed a continuing concern for meeting individual differences, sex differences in cognitive modes of thinking have largely been ignored (2).

In almost all studies reviewed, underachieving boys outnumbered underachieving girls. About two-thirds of school retentions are boys (8). Educators have suggested several reasons for this. In addition to the boys' maturational lag, and the use of methods and materials more suitable for girls' learning style, mentioned previously, teacher attitude and curriculum content may be significant factors. Boys are aware that teachers seem to prefer girls in the classroom (7). Some educators have suggested that different emphases should be placed on subjects like biology, geography, history, and language for boys and girls (2).

Reviewing a wide bibliography of physiological and psychological studies, Bentzen concluded that

> *. . . among school-age children, boys tend to be biophysically less mature than girls of the same chronological age, and that, finally, sex ratios reported in studies of a wide range of learning and behavior disorders include a significantly higher proportion of males than females (1).*

While testing instruments presently available show no difference between the sexes in basic intelligence, achievement tests generally show girls superior in language skills, boys in arithmetic reasoning, and science (9). A review of the literature shows these results to be consistent, and some studies are quite impressive.

There is also substantial evidence indicating that boys receive lower grades than girls. Kremer reviewed thirty-five studies undertaken between 1935–1965 and reported this to be true (5). Tyler summarized another group of studies and concluded that *"the evidence from numerous studies of sex differences in school achievement is remarkably consistent in one respect: girls are assigned higher grades by their teachers than boys are"* (10). Grambs noted that *"it is not right to deny girls a full development of their intellectual powers by keeping them to the lower level of the more slowly-maturing boys"* (2). Mead has described the unwholesome habits *"many girls develop, especially in early adolescence, of downgrading their own academic efforts in order to be popular with and accepted by boys"* (6).

Problem

Are academic achievement, self-discipline, self-concept, sex role identification, and attitude toward school different for seventh and eighth grade boys and girls who are enrolled in classes of the same sex, as opposed to those grouped together? Five null hypotheses were derived from these variables and tested at the .05 level of confidence.

Procedures

The investigation adhered to the pretest posttest control group experimental design with groups randomly selected. Two junior high schools in northern Illinois cooperated in the study. Students in each school (N = 150) were randomly selected for the study.[2] Classes were same-sex organized at the seventh and eighth grade levels in the areas of science, mathematics, social studies, and English. For practical purposes, sex segregation was achieved in physical education and home economics. The students spent approximately five-sixths of the school day for one school year in this organizational pattern.

Controls. Course content, class size, instructional materials, general teaching style, and time involved were similar for all students in the study. The experimental treatment was constant for all experimental groups. Application of analysis of covariance in treatment of the data provided additional statistical controls. Furthermore, data gathered from the students by interviews were obtained by a member of the research staff of the same sex as the student. No control was provided for teacher associated factors.

Data. In addition to measurement of the dependent variables using the instruments of the study, data relating to each dependent variable were obtained from the students' records and from the teachers and counselors who were responsible for these students.

Variables and Experimental Treatment. The independent variable was the same-sex class organization for seventh and eighth grade students and operated over a period of one school year (9 months). The treatment involved classes in English, mathematics, social studies, science, and physical education, or approximately five-sixths of the students' school day. Dependent variables were academic achievement, self-discipline, self-concept report, sex-role identification, and attitudes toward school.

Instruments. Dependent variables were measured by pretests and posttests in September and May, using the Iowa Test of Basic Skills, Metropolitan Achievement Tests, Standard Achievement Test, California Test of Personality, and the Ellis-Many-Frey Report of Self-Concept Scale. Additional data were gathered by interviews and teacher reports.

Analysis of Data. Because of the differences in the nature of the student bodies, each of the two schools cooperating in the study was treated separately, thereby having the effect of two studies. Statistical treatment included a comparison of the differences between the means of the experimental and control groups by application of analysis of covariance. More specifically, a

2 × 2 × 2 factorial design was used to enable comparisons of the main effects of class grouping, sex, grade level, and the interaction effects of the independent variable upon the criterion measures to be made. Each of the five null hypotheses was tested in this manner. Attention was directed to change differences for groups as well as to differences between and among group means.

Results

Hypothesis 1. There is no significant difference in academic achievement between a group of junior high school boys and girls after one year in classes of the same sex.

Pretest and posttest data from the standardized measures of achievement and the marks which teachers assigned to students for the areas where the experimental treatment prevailed were treated statistically by analyses of covariances for each of the two schools cooperating in the study. Statistically significant differences attributable to the experimental treatment were not found; thus, the hypothesis was supported at both schools (see Tables 1, 2, and 3).

It should be noted that an apparent statistically significant difference was found (See Table 1) regarding the I × J × K interaction for composite achievement scores. However, the treatment of the individual components of achievement failed to yield a significant difference. Thus, it was concluded that the interaction was inconsistent and spurious and did not warrant rejection of the hypothesis. (Space precludes the tabular presentation of each of the thirty-five analyses of covariance performed in the testing of Hypothesis 1.)

Table 1.—Summary of Analysis of Covariance for Iowa Tests of Basic Skills (Composite Score) Adjusted Posttest Scores (School A)

Source	NDF	SS	Mean SQ
I Effect (Sex)	1.0000	23.4766	23.4766
J Effect (Grade)	1.0000	6.5039	6.5039
K Effect (Group)	1.0000	0.0039	0.0039
I x J	1.0000	22.4883	22.4883
I x K	1.0000	1.8555	1.8555
J x K	1.0000	21.4922	21.4922
I x J x K	1.0000	169.0781	169.0781*
Within	79.0000	2039.4570	25.8159
TOTAL	86.0000	2284.3555	

*F=6.5493, significant at the .05 level; F required at .01=7.00 and at .05=3.97.

Table 2.—Summary of Analysis of Variance for GPA's Adjusted Posttest Scores (School A)**

Source	NDF	SS	Mean SQ
K Effect (Group)	1	13.4233	13.4233*
I Effect (Sex)	1	2.5066	2.5066
I x J	1	0.0222	0.0222
Within	45	94.0596	2.0902
TOTAL	48	110.0117	

*F=6.4420, F required at .01=6.93 and at .05=3.95.

**Grades for eighth grade students only.

Hypothesis 2. There is no significant difference

Table 3.—Summary of Analysis of Covariance for GPA's Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	25.3967	25.3967
K Effect (Group)	1	11.2271	11.2271
I Effect (Sex)	1	41.5605	41.5605*
I x J	1	10.7126	10.7126
I x K	1	1.2485	1.2485
J x K	1	8.2070	8.2070
I x J x K	1	0.3835	0.3835
Within	107	394.2605	3.6847
TOTAL	114	492.9966	

*F=11.2792, F required at .01=6.93 and at .05=3.95.

in self-discipline between a group of junior high school boys and girls after one school year in classes of the same sex.

Pretest and posttest data from the "Self-Reliance" scale of the California Test of Personality and Tasks 8, "Desiring and Achieving Socially Responsible Behavior," and 11, "Developing Conscience, Morality and a Scale of Values to Guide Behavior," from the Ellis-Many-Frey Report of Self-Concept scale, were treated statistically by analyses of covariance for each of the two schools. Statistically significant differences attributable to the experimental treatment were not found; thus, the hypothesis was supported (See Tables 4, 5, 6, and 7).

Table 4.—Summary of Analysis of Covariance for Task 11, "Developing Conscience, Morality, and a Scale of Values to Guide Behavior" Adjust Posttest Scores (School A)

Source	NDF	SS	Mean SQ
I Effect (Sex)	1	79380.7500	79380.7500*
J Effect (Grade)	1	2887.6875	2887.6875
K Effect (Group)	1	1957.9375	1957.9375
I x J	1	8862.6875	8862.6875
I x K	1	5944.3750	5944.3750
J x K	1	322.8125	322.8125
I x J x K	1	22818.8125	22818.8125
Within	79	713408.3750	9030.4844
TOTAL	86	835583.4375	

*F=8.7903, significant at .01 level; F required at .01=7.00 and at .05=3.97.

Table 5.—Summary of Analysis of Covariance for Task 11, Adjust Posttest Scores (School B)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	994.0000	994.0000
K Effect (Group)	1	52726.0000	52726.0000
I Effect (Sex)	1	25356.0000	25356.0000
I x J	1	228.0000	228.0000
I x K	1	10573.0000	10573.0000
J x K	1	926.0000	926.0000
I x J x K	1	467.0000	467.0000
Within	103	1412029.0000	13709.0156
TOTAL	110	1503299.0000	

*All F's nonsignificant; F required at .01=6.93 and at .05=3.95.

Table 6.—Summary of Analysis of Covariance for "Self-Reliance" Scale, California Test of Personality Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
I Effect (Sex)	1	14.6631	14.6631
J Effect (Grade)	1	2.4961	2.4961
K Effect (Group)	1	1.6404	1.6404
I x J	1	0.7512	0.7512
I x K	1	0.2246	0.2246
J x K	1	0.3557	0.3557
I x J x K	1	0.0308	0.0308
Within	79	375.2866	4.7505
TOTAL	86	395.4485	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97

Hypothesis 3: There is no significant difference in report of self-concept of ability between a group of junior high school boys and girls after one school year in classes of the same sex.

The California Test of Personality (CTP) was used to test the hypothesis. For purposes of this research, one individual scale, "Sense of Personal Worth," and the three composite scores, "Total Personal Adjustment," "Total Social Adjustment," and the "Total Composite Scores," of all fourteen sub-scores were applied. The data presented in Tables 8–15 indicate that no statistically significant differences were found that could be attributed to the experimental treatment; therefore, the hypothesis was supported.

Table 7.—Summary of Analysis of Covariance for "Self-Reliance" Scale, California Test of Personality Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	1.1802	1.1802
K Effect (Group)	1	0.7849	0.7849
I Effect (Sex)	1	4.2253	4.2253
J x K	1	0.9131	0.9131
J x I	1	37.4114	37.4114*
K x I	1	1.0752	1.0752
J x K x I	1	0.3423	0.3423
Within	103	513.8953	4.9893
TOTAL	110	559.8276	

*F=7.4983, significant at .01 level; F required at .01=6.93 and at .05=3.95.

Table 8.—Summary of Analysis of Covariance for "Sense of Personal Worth" Scale, CTP Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	13.8899	13.8899
I Effect (Sex)	1	0.4956	0.4956
K Effect (Group)	1	4.4878	4.4878
J x I	1	2.6697	2.6697
J x K	1	5.6880	5.6880
I x K	1	0.5076	0.5076
J x I x K	1	10.0305	10.0305
Within	79	469.7114	5.9457
TOTAL	86	507.4805	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 9.—Summary of Analysis of Covariance For "Total Personal Adjustment," CTP Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	292.9609	292.9609
I Effect (Sex)	1	22.8008	22.8008
K Effect (Group)	1	45.0234	45.0234
J x I	1	9.4883	9.4883
J x K	1	139.9492	139.9492
I x K	1	10.8086	10.8086
J x I x K	1	4.3516	4.3516
Within	79	7027.7031	88.9583
TOTAL	86	7553.0859	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 10.—Summary of Analysis of Covariance For Total Social Adjustment, CTP Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	0.6445	0.6445
I Effect (Sex)	1	14.9727	14.9727
K Effect (Group)	1	165.9062	165.9062
J x I	1	2.0547	2.0547
J x K	1	257.8555	257.8555
I x K	1	27.4805	27.4805
J x I x K	1	7.4609	7.4609
Within	79	8435.7109	106.7811
TOTAL	86	8912.0859	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

able 11.—Summary of Analysis of Covariance For Total djustment, CTP Adjusted Posttest Scores (School A)*

ource	NDF	SS	Mean SQ
Effect (Grade)	1	404.4883	404.4883
Effect (Sex)	1	116.6289	116.6289
Effect (Group)	1	18.8047	18.8047
x I	1	28.8906	28.8906
x K	1	771.1680	771.1680
x K	1	83.7656	83.7656
x I x K	1	39.1562	39.1562
ithin	79	25461.0078	322.2910
OTAL	86	26923.9102	

All F's nonsignificant; F required at 01=7.00 and at .05=3.97.

ble 12.—Summary Analysis of Covariance For "Sense Personal Worth" Scale, CTP Adjusted Posttest Scores chool B)*

ource	NDF	SS	Mean SQ
Effect (Grade)	1	7.2241	7.2241
Effect (Sex)	1	17.7280	17.7280
Effect (Group)	1	12.9446	12.9446
x I	1	4.7036	4.7036
x K	1	21.9275	21.9275
x K	1	13.9824	13.9824
x I x K	1	26.6282	26.6282
ithin	103	960.9351	9.3295
OTAL	110	1066.0735	

All F's nonsignificant, F required at 01=6.93 and at .05=3.95.

Hypothesis 4: There is no significant difference in sex role identification between a group of junior high school boys and girls after one school year in classes of the same sex.

Table 13.—Summary Analysis of Covariance For Total Personal Adjustment, CTP Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	72.9492	72.9492
I Effect (Sex)	1	176.5937	176.5937
K Effect (Group)	1	68.5625	68.5625
J x I	1	46.3164	46.3164
J x K	1	620.8594	620.8594*
I x K	1	268.4883	268.4883
J x I x K	1	10.3750	10.3750
Within	103	11540.0352	112.0392
TOTAL	110	12804.1797	

*F=5.5414, significant at .05 level; F required at .01=6.93 and at .05=3.95.

Table 14.—Summary of Analysis of Covariance For Total Social Adjustment, CTP Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	701.4922	701.4922*
I Effect (Sex)	1	229.8828	229.8828
K Effect (Group)	1	346.0859	346.0859
J x I	1	15.3047	15.3047
J x K	1	92.0117	92.0117
I x K	1	83.2539	83.2539
J x I x K	1	6.1172	6.1172
Within	103	10000.6641	97.0938
TOTAL	110	11474.8125	

*F=7.2248, significant at .01 level; F required at .01=6.93 and at .05=3.95.

Table 15.—Summary of Analysis of Covariance For Total Adjustment, CTP Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	1310.2070	1310.2070[a]
I Effect (Sex)	1	771.7852	771.7852
K Effect (Group)	1	99.6992	99.6992
J x I	1	144.4961	144.4961
J x K	1	1216.7930	1216.7930[b]
I x K	1	695.9102	696.9102
J x I x K	1	15.4609	15.4609
Within	103	27706.9414	268.9993
TOTAL	110	31961.2930	

[a] significant at .05 level, F=4.8706;
[b] significant at .05 level, F=4.5234;
F required at .01=6.93 and at .05=3.95.

Table 16.—Summary of Analysis of Covariance For Task 2, E-M-F Report of Self-Concept Scale Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
I Effect (Sex)	1	9945.6875	9945.6875
J Effect (Grade)	1	5121.5000	5121.5000
K Effect (Group)	1	5834.8750	5834.8750
I x J	1	14433.6875	14433.6875
I x K	1	5927.6875	5927.6875
J x K	1	9502.3750	9502.3750
I x J x K	1	109.8125	109.8125
Within	79	727652.5000	9210.7891
TOTAL	86	778528.1250	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 17.—Summary of Analysis of Covariance For Task 2, E-M-F Report of Self-Concept Scale (Achieving A Masculine or Feminine Social Role) Adjusted Posttest Scores (School A)*

Source	SS	DF	MS
I Effect (Sex)	11394.2500	1	11394.2500
J Effect (Grade)	2259.3125	1	2259.3125
K Effect (Group)	1510.6250	1	1510.6250
I x J	3900.1250	1	3900.1250
I x K	8879.8750	1	8879.8750
J x K	28745.5000	1	28745.5000
I x J x K	70.1875	1	70.1875
Within	664190.8125	79	8407.4766
TOTAL	720950.6875	86	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 18.—Summary of Analysis of Covariance For Task 3, E-M-F Report of Self-Concept Scale (Building Wholesome Attitudes Toward Oneself as a Growing Organism) Adjusted Posttest Scores (School A)

Source	SS	DF	MS
I Effect (Sex)	31485.6250	1	31485.6250*
J Effect (Grade)	5162.5625	1	5162.5625
K Effect (Group)	1985.5625	1	1985.5625
I x J	470.0000	1	470.0000
I x K	232.8125	1	232.8125
J x K	2076.4375	1	2076.4375
I x J x K	628.5000	1	628.5000
Within	623809.3125	79	7896.3203
TOTAL	665850.8125	86	

*F=3.9873, significant at .05 level; F required at .01=7.00 and at .05=3.97, The three measures selected to test this hypothesis were Tasks 1, 2, and 3 of Form J of the Ellis-Many-Frey (E-M-F) Report of Self-Concept Scale. Task 1 asked pupils to rate themselves on in-school and out-of-school activities to indicate how well they felt they were achieving new and more mature relationships with age-mates of both sexes. They evaluated aspects of their growth and behavior that were helping them develop an appropriate masculine or feminine sex role in Task 2. Task 3 asked pupils to report their feelings of acceptance about their growing bodies. Items were adjusted for boys and girls.

Data reported in Tables 16 through 21 show that no significant differences were found which were associated with the effect of same sex class grouping and the hypothesis was supported.

Table 19.—Summary of Analysis of Covariance For Task 2, E-M-F Report of Self-Concept Scale Adjusted Posttest Scores (School B)*

Source	NDF	SS	Mean SQ
I Effect (Sex)	1	22459.0000	22459.0000
J Effect (Grade)	1	3927.0000	3927.0000
K Effect (Group)	1	16331.0000	16331.0000
I x J	1	1614.0000	1614.0000
I x K	1	16037.0000	16037.0000
J x K	1	921.0000	921.0000
I x J x K	1	4762.0000	4762.0000
Within	103	1094843.0000	10629.5430
TOTAL	110	1160894.0000	

*All F's nonsignificant, F required at .01=6.93 and at .05=3.95.

Hypothesis 5: There is no significant difference in attitude toward school between a group of junior high school boys and girls after one year in classes of the same sex.

Four sets of data were used to test this hypothesis: the "School Relations Scale" of the CTP, Task 9, "Developing Attitudes Toward Groups and Social Institutions"; Task 10, "Developing Intellectual Skills and Concepts Necessary for Civic Competence," of the E-M-F Report of Self-Concept Scale; and student responses to interviews.

Table 20.—Summary of Analysis of Covariance For Task 2, E-M-F Report of Self-Concept Scale Adjusted Posttest Scores (School B)*

Source	SS	DF	MS
I Effect (Sex)	14596.0000	1	14596.0000
J Effect (Grade)	11641.0000	1	11641.0000
K Effect (Group)	1631.0000	1	1631.0000
I x J	6263.0000	1	6263.0000
I x K	7353.0000	1	7353.0000
J x K	4197.0000	1	4197.0000
I x J x K	6270.0000	1	6270.0000
Within	1287838.0000	103	12503.2812
TOTAL	1339789.0000	110	

*All F's nonsignificant; F required at .01=6.93 and at .05=3.95.

Table 21.—Summary of Analysis of Covariance For Task 3, E-M-F Report of Self-Concept Scale Adjusted Posttest Scores (School B)*

Source	NDF	SS	Mean SQ
I Effect (Sex)	1	40.0000	40.0000
J Effect (Grade)	1	364.0000	364.0000
K Effect (Group)	1	10395.0000	10395.0000
I x J	1	1368.0000	1368.0000
I x K	1	4703.0000	4703.0000
J x K	1	37942.0000	37942.0000
I x J x K	1	24228.0000	24228.0000
Within	103	1244432.0000	24228.0000
TOTAL	110	1323472.0000	12081.8633

*All F's nonsignificant; F required at .01=6.93 and at .05=3.95.

No significant differences were found attributable to grouping and, again, the null hypothesis was supported (See Tables 22–26).

Table 22.—Summary of Analysis of Covariance For Task 9, E-M-F Report of Self-Concept Scale (Developing Attitudes Toward Social Groups and Institutions) Adjusted Posttest Scores (School A)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	36379.6250	36379.6250
I Effect (Sex)	1	922.4375	922.4375
K Effect (Group)	1	3085.1250	3085.1250
J x I	1	901.0000	901.0000
J x K	1	12353.0625	12353.0625
I x K	1	3102.0625	3102.0625
J x I x K	1	15282.3750	15282.3750
Within	79	847762.6875	10731.1719
TOTAL	86	919788.3750	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 23.—Summary of Analysis of Covariance For Task 9, E-M-F Report of Self-Concept Scale Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	1007.0000	1007.0000
I Effect (Sex)	1	89845.0000	89845.0000*
K Effect (Group)	1	48440.0000	48440.0000
J x I	1	448.0000	448.0000
J x K	1	21351.0000	21351.0000
I x K	1	19320.0000	19320.0000
J x I x K	1	9938.0000	9938.0000
Within	103	1573863.0000	15280.2227
TOTAL	110	1764212.0000	

*F=5.8798, significant at .05 level; F required at .01=6.93 and at .05=3.95.

General Observations. While statistically significant differences were lacking regarding the effect of same-sex class grouping on the dependent variables of the study, several observations were made:

1. *Girls tended to receive higher teacher assigned marks then did boys (see Tables 2 and 3).*

Table 24.—Summary of Analysis of Covariance For "School Relations" Scale, CTP Adjusted Posttest Scores (School B)*

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	7.1523	7.1523
I Effect (Sex)	1	4.1624	4.1624
K Effect (Group)	1	0.4602	0.4602
J x I	1	3.9978	3.9978
J x K	1	12.7573	12.7573
I x K	1	9.5532	9.5532
J x I x K	1	12.8237	12.8237
Within	79	530.1863	6.7112
TOTAL	86	581.0933	

*All F's nonsignificant; F required at .01=7.00 and at .05=3.97.

Table 25.—Summary of Analysis of Covariance For "School Relations" Scale, CTP Adjusted Posttest Scores (School B)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	16.0317	16.0317
I Effect (Sex)	1	11.2334	11.2334
K Effect (Group)	1	71.4519	71.4519*
J x I	1	1.3464	1.3464
J x K	1	5.0977	5.0977
I x K	1	1.3425	1.3425
J x I x K	1	4.0498	4.0498
Within	103	705.1418	6.8460
TOTAL	110	815.6953	

*F=10.4370, significant at .01 level; F required at .01=6.93 and at .05=3.95.

Table 26.—Summary of Analysis of Covariance For Answers To Interviewer's Question, "How Do You Like School This Year?" Adjusted Posttest Scores (School A)

Source	NDF	SS	Mean SQ
J Effect (Grade)	1	1.1798	1.1798
I Effect (Sex)	1	14.7632	14.7632*
K Effect (Group)	1	3.0180	3.0180
J x I	1	0.0359	0.0359
J x K	1	2.3780	2.3780
I x K	1	2.7793	2.7793
J x I x K	1	2.5219	2.5219
Within	79	73.7355	0.9334
TOTAL	86	100.4118	

*F=15.8165, significant at .01 level; F required at .01=7.00 and at .05=3.97.

2. *Girls tended to score higher on standardized achievement tests than did boys.*
3. *Eighth grade students possessed a more wholesome attitude toward themselves than did seventh graders (see Table 18).*
4. *Girls reported attitudes that were more favorable toward social groups and institutions than did boys (see Table 23).*
5. *Boys reported a more favorable attitude toward school than did girls under conditions of this experiment (see Tables 25 and 26). Thus, the reactions to experience in the same-sex grouping pattern was less negative for boys than for girls.*
6. *The attitudes of students and teachers toward same-sex grouping became more negative as the year progressed.*

Summary, Conclusions, and Recommendations

The study sought to determine the effect of same-sex class organization on junior high school students' academic achievement, self-discipline, self-concept, sex role identification, and attitude toward school. Five null hypotheses (one for each dependent variable) were tested by sixty-one analyses of covariance.

Findings failed to yield significant differences which were attributable to same sex grouping; thus, all hypotheses of the study were supported. It was concluded that same-sex class organization for seventh and eighth grade students, so grouped for approximately four-fifths of the school day for the school year of 9 months, had no effect regarding the dependent variables of the study.

It should be noted that girls tended to have received higher marks from teachers, achieved higher scores on achievement tests, reported more favorable attitudes toward social groups and institutions, and responded more negatively to the sex-segregated school experience than did boys. These findings generally were consistent with the literature about early adolescents.

It is recommended that additional research in this area be conducted over a longer period of time, controlled for the teacher factor, and provide a higher degree of student sex segregation than was present in the investigation reported here.

FOOTNOTES

1. This study was supported by a grant from the U.S. Office of Education (Project No. 7-E115, Grant No. OEG-08-00015-0218(010).
2. Due to class scheduling for music, student council, etc., some adjustments had to be made.

REFERENCES

1. Bentzen, Frances, "Sex Ratios in Learning and Behavior Disorders," *National Elementary Principal,* 46:13–17, November 1966.
2. Grambs, Jean D.; Waetjen, Walter B., "Being Equally Different," *National Elementary Principal,* 46: (no. 2), November 1966; "Sex Differences: A Case of Educational Evasion?" *Teachers College Record,* 65: 261–271, December 1963.
3. Havighurst, Robert J., *Developmental Tasks and Education,* University of Chicago Press, Chicago, Illinois, 1948.
4. Kolesnik, Walter B., "Sex Differences and Education," *America,* 108:552–555, April 1963.
5. Kremer, Bruce J., "Is Coeducation Unfair to Boys?" *Catholic School Journal,* 65:37–39, October 1965.
6. Mead, Margaret, "The Early Adolescent in Today's American Culture and Implications for Education," *Junior High School Newsletter,* 1:1–6, February 1963.
7. Meyer, William J.; Thompson, George G., "Sex Differences in the Distribution of Teacher Approval and Disapproval Among Sixth-Grade Children," *Journal of Educational Psychology,* 47:361–368, December 1965.
8. Teigland, John J.; Winkler, Ronald C., "Is Underachievement Basically a Male Problem?" *Personnel and Guidance Journal,* 44:430–431, December 1955.
9. Terman, Lewis M.; Tyler, Leona E., "Psychological Sex Differences," in Carmichael, Leonard (ed.), *Manual of Child Psychology,* John Wiley and Sons, Inc., New York, 1954, pp. 1064–1114.
10. Tyler, F., "Individual and Sex Differences," in Harris, Chester W. (ed.), *Encyclopedia of Educational Research.* Third Edition, The Macmillan Company, New York, 1960, pp. 680–688.
11. Woody, Thomas, *A History of Women's Education in the United States,* Vol. II, 2 vols., The Science Press, New York, 1929.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Nature and Language of Scientific Terms In the Arabic Press of the Middle East[1]

WADI' D. HADDAD, American University of Beirut

ABSTRACT

This study was designed to determine the relationship between the disciplinary nature of scientific terms used in the Arabic press and their linguistic structure, as well as the association between these attributes and the theme of articles and the countries involved. The study also sought to determine the degree of consistency in the usage of such terms and the difference between press and authorized terms. Technical terms found in scientific articles appearing in monthly issues of eight magazines and a selected sample of newspapers of eight countries over a period of 18 months were recorded, classified, and analyzed. Results showed that the reading of scientific articles in the Arabic press requires a considerable amount of technical knowledge in the various disciplines of science. χ^2 tests ($\alpha = 0.05$) revealed a significant association between the nature of terms and the theme of articles as well as the country in which they were used, but the degree of association, measured by ϕ', was weak. A high proportion of these terms was either totally or partially borrowed, and their distribution according to their linguistic structure depended upon the discipline but not upon the country. The problem of borrowing and transliteration led to inconsistencies in the usage of terms. It was also found that journalists did not deviate from authorized terms, when available, and that, in one-fourth of the cases, they had to use terms that have not yet been authorized.

SCIENCE IN the twentieth century is one of the most prolific disciplines. The expanding dimensions of scientific knowledge have greatly accelerated the rate of technological achievements. The activities as well as the products of science are becoming an integral part of our society. The average informed citizen follows with excitement the Apollo flights, marvels at the operations of the transplantation of hearts, worries about the missile race, and discusses the implications of artificial insemination.

Mass media are playing a significant role in making information on scientific developments available. As a result of a study conducted in 1959, Withey (11) reported that contact of the adult public with science in the U.S. through mass media (i.e., newspapers, magazines, radio, and television) was much broader than their contact through regular science courses. Two out of three could recall science news items which they encountered. *"Further inquiry clearly showed that newspapers are the major carriers. TV and magazines run about an equal second"* (11:155). Other studies related to science content of newspapers and magazines revealed that a considerable number of articles dealt with science and/or technology (4, 6, 8, 9, 10).

Newspapers and magazines in the Arab Middle East are no exception to this phenomenon.[2] Since these media are an essential source of scientific information it is important to find out the level of prerequisite knowledge that the reader must have in order to understand the science articles in the press. A study of this prerequisite knowledge involves the examination of the nature of scientific terms and phrases used in these articles. On the other hand, since the Arab Middle East is not a "producing" area of scientific development, its scientific literature is, directly or indirectly, a translation from English or French. This operation poses the problem of the kind of terms that should be used to express new concepts.

To cope with this problem many of the Arab countries, particularly Egypt, Iraq, Syria, and Morocco, have established academies of Arabic language. One of the functions of these academies is

> *to make the Arabic language meet the demands of the progress of the sciences and the arts . . . by specifying, by means of dictionaries, special commentaries, or other methods, the terms and structures that should be used and the ones that should be avoided* (1:4).

The procedure followed in the selection of scientific terms is best expressed by the following policy statement made by the Academy of Arabic Language in Cairo:

> *The time has passed for members of the Academy to invent scientific technical terms or to select for them words that are obsolete or extraneous. They now believe that their primary duty is to record terms that are established by tradition and supported by usage* (1:6).

It is obvious that many new developments in science, when covered by the press, require new or restructured words. These terms are put into use and consequently may be established later by academies of the Arabic language. In such a case the scientific terminology of newspapers and magazines can be a major source of the established scientific vocabulary in the Arab Middle East. Consequently, the study of scientific terms in the Arab Middle Eastern press can have a revealing as well as a predictive value in the understanding of the relationship between scientific terminology and the Arabic language.

Purpose of the Study

The aim of this study is to investigate the nature of scientific terms found in and prerequisite to the reading of articles dealing with science and/or technology in a selected number of newspapers and magazines in the Middle East by determining the following:

1. Distribution of the scientific terms when a classification system based on the conventional disciplinary lines of biology, chemistry, etc., is imposed, and the relationship between this distribution and (a) the theme of articles and (b) the country in which newspapers were published.

2. Distribution of the scientific terms according to a classification based on the linguistic structure of these terms, and the relationship between this classification and (a) the classification based on the disciplinary scheme and (b) the country in which newspapers were published.

3. The difference between the scientific terminology established by the Academy of Arabic Language in Cairo and the terminology used by the Egyptian press.

4. The degree of consistency in the usage of scientific terms.

Procedure

Seven magazines that appeal to the general public and the circulation of which is not limited to one Arab country were selected. To these was added a family magazine which circulates widely among women. One monthly issue of each of these eight magazines between January 1967 and June 1968 was selected for analysis. The selected magazines were: 1. *El-'Arabi,* 2. *El-Jadeed,* 3. *El-Khartoum,* 4. *El-Musawwar,* 5. *Rose el-Yusuf,* 6. *Es-Sayyaad,* 7. *El-Usbu' el-'Arabi,* and 8. *Usrati.*

In addition to these magazines, three capital newspapers were selected from each of eight Arab countries on the basis of high circulation and representation. Nine dates were randomly selected for each month separately and distributed randomly and equally over the three newspapers. In cases where only two papers are published the dates for the third paper were alternately divided between them. Issues selected in this manner from January 1967 to June 1968 were read for analysis. Table 1 shows the countries chosen as well as the newspapers read and analyzed.

The selected magazines and newspapers were thoroughly read, and all articles dealing with science and/or technology were isolated. According to its theme, each article was then forced into one of the following categories:

a. Life science including zoology, medicine and health.

b. Agricultural science including botany and agricultural products.

c. Mechanical science including equipment, instruments, machines, and computers.

d. Space science including astronomy, equipment for space flights and rocketry.

e. Earth science including natural resources, minerals, petroleum industry, and geology.

f. Matter and Energy including the structure of matter, chemistry, and nuclear energy.

Technical terms that appeared in the selected articles without being defined or explained were considered prerequisite to the understanding of these articles. These terms were recorded and forced into one of the conventional branches of science (see Table 2). Each term was also classified linguistically according to the following scheme:

Type A: are loan-words that are transliterated without any structural modifications. For example, *Oxygen; microscope; radar.*

Type B: are loan words that are modified according to the rules of derivation in Arabic. For example, *lymphawi* (lymphatic); *aksada* (meaning oxidized and derived from oxide); *electronaat* (plural of electron); *balmarat* (meaning polymerization and derived from polymer).

Type C: are scientific terms that are genuinely Arabic.

Table 1.—Listing of Countries and Newspapers

Country	Newspaper 1	Newspaper 2	Newspaper 3
1. Egypt	El-Akhbar	El-Ahram	El-Jumhuriyya
2. Iraq	Eth-Thawra	El-Jumhuriyya	
3. Jordan	Ed-Dustur	Ed-Difaʻ	El-Jihad
4. Kuwait	Er-Raʼyo el- Aam*		
5. Lebanon	En-Nahar	Lisan el-Ḥaal	El-Anwar
6. Saudi Arabia	El-Madina	El-Bilaad	Er-Riyad
7. Sudan	Er-Raʼyo el- Aam**	Es-Sudan el-Jadid	El-Ayyam
8. Syria	Eth-Thawra	El-Baʻth	

* Issues of this newspaper only were available. Three issues per month were selected.
** Issues of only 6 months (January-June 1967) were available.

Type D: are scientific terms that are expressed in Arabic with equivalent in a foreign language. For example, *fitr (fungi); ʻamil musaʻid (catalyst); inzilam el-ʻadasa (cataract).*

Type AC: are composite terms combining types A and C. For example, *thani okseed el-carboon* (carbon dioxide); *iltihab el-bariton* (peritonitis); *Hamid en-nitreek* (nitric acid).

Type BC: are composite terms combining types B and C. For example: *murakkabat keemawiyya* (chemical compounds); *mijhar electroni* (electronic microscope); *tawazon hormoni* (hormonal balance).

To test for statistical association between categorical attributes (disciplines, countries, etc.), the Pearson chi-square test was employed with an $\alpha = 0.5$ level of significance. The strength of association, however, is described by Cramer's

Table 2.—Number and Percentage of Items of Prerequisite Knowledge Classified According to Disciplines

Science Discipline	Magazines No.	Magazines %	Newspapers No.	Newspapers %
Astronomy	23	5.2	69	6.8
Biology	224	50.8	418	41.5
Chemistry	93	21.1	183	18.2
Engineering	12	2.7	85	8.4
Geology	3	.7	37	3.7
Physics	86	19.5	216	21.4
Total	441	100.0	1008	100.0

Table 3.—A Breakdown of Biology Terms Into Subcategories

Subcategories	Magazines No.	Magazines %	Newspapers No.	Newspapers %
Anatomy	59	26.3	111	26.7
Physiology	14	6.3	34	8.2
Diseases and Disorders	62	27.7	128	30.8
Diagnosis and Treatment	13	5.8	44	10.6
Genetics and Developmental Biology	24	10.7	19	4.6
Taxonomy and Ecology	29	12.9	41	9.9
Biochemistry	23	10.3	38	9.2
Total	224	100.0	415*	100.0

*The discrepancy between this figure and the one appearing in Table 2 is due to the fact that some biology items were too general to be classified into the subcategories. Other discrepancies appearing in later tables have the same reason.

Table 4.—A Breakdown of Chemistry Terms Into Subcategories

Subcategories	Magazines No.	Magazines %	Newspapers No.	Newspapers %
Inorganic	37	39.8	58	32.0
Organic	18	19.3	28	15.5
Chemicals	38	40.9	95	52.5

statistic ϕ' which is defined as

$$\phi' = \sqrt{\frac{\chi^2}{N\,(L-1)}}$$

where N is the number of observations and L is the smaller of the number of rows or the number of columns in a χ^2 test of association (3).

Findings

Disciplinary Nature of Scientific Terms. It is noted from Table 2 that there are 441 items of prerequisite knowledge that are essential to the understanding of scientifically or technologically oriented magazine articles, and 1,008 items prerequisite to the understanding of newspaper articles. About half of the terms are drawn from biology, one-fifth from physics and another fifth from chemistry. Very little is drawn from the other disciplines. Of the biology terms, about half are related to anatomy and diseases and disorders, while there is a marked scarcity of physiological terms (see Table 3). The chemistry terms are about equally divided between names of chemicals and conceptual terms. Of the latter category, there are about twice as many terms related to inorganic chemistry as to organic chemistry (see Table 4). The physics terms are drawn mainly from electricity and magnetism and nuclear and atomic physics. Very few terms deal with heat (see Table 5).

Table 5.—A Breakdown of Physics Terms Into Subcategories

Subcategories	Magazines No.	Magazines %	Newspapers No.	Newspapers %
Electricity and Magnetism	27	31.4	62	29.0
Heat	3	3.5	18	8.4
Mechanics	10	11.6	34	15.9
Nuclear and Atomic Physics	30	34.9	60	28.0
Optics and Waves	16	18.6	40	18.7

Chi-square tests of association revealed that there is a significant difference between the distributions of magazine terms and newspaper terms according to the disciplinary scheme,

Table 6.—Distribution of Scientific Terms According to Pella's Study

Discipline	Magazines No.	Magazines %	Newspapers No.	Newspapers %
Astronomy	4	2.2	45	9.7
Biology	135	73.8	224	48.4
Anatomy	18		24	
Physiology	18		25	
Diseases and Disorders	48		74	
Diagnosis and Treatment	26		40	
Genetics and Developmental Biology	16		16	
Taxonomy and Ecology	2		33	
Biochemistry	7		12	
Chemistry	10	5.5	38	8.2
Inorganic	7		25	
Organic	0		4	
Chemicals	3		9	
Engineering	3	1.6	7	1.5
Geology	1	.5	19	4.1
Physics	30	16.4	130	28.1
Electricity and Magnetism	4		32	
Heat	0		11	
Mechanics	13		43	
Nuclear and Atomic Physics	9		30	
Optics and Waves	4		14	
Total	183		463	

($\chi^2 = 34.12$, df = 5). The strength of association between the two categorical attributes, however, is low ($\phi' = 0.15$). There is also a significant difference between the two distributions of biology terms ($\chi^2 = 14.56$, df = 6) with a weak degree of association ($\phi' = 0.15$). On the other hand, no significant difference exists between the chemistry ($\chi^2 = 3.32$, df = 2) or the physics distributions ($\chi^2 = 3.96$, df = 4).

For the purpose of comparison, the scientific terms reported by Pella and others (8) in a similar study were re-classified according to the scheme used in this study (see Table 6). Pella's study was based on the analyses of twenty-two capital city newspapers and eight magazines published in the U.S. over a period of 6 months. When the distributions of the scientific terms in the Middle Eastern Arabic press and the American press were compared it was found that the two distributions (according to the major discipline categories) differ significantly, both for the magazines ($\chi^2 = 35.07$, df = 5) and the newspapers ($\chi^2 = 58.16$, df = 5). In particular, the American magazines had a higher proportion of biology terms and a lower proportion of chemistry and physics terms, and the American newspapers had a higher proportion of biology terms and a lower one of chemistry and engineering terms.

When the distributions of terms into sub-categories within each of the major disciplines of biology, chemistry, and physics were compared for the two studies it was found that each pair of respective distributions differs significantly for both newspapers and magazines except for chemistry where there was no significant difference between the two distributions of terms appearing in magazines. In specific, the proportions of terms dealing with anatomy, biochemistry, organic chemistry, chemicals, electricity and magnetism, atomic and nuclear physics, and optics and waves were higher in the Middle Eastern press, while the American press had high proportions of terms expressing concepts of physiology, diseases, treatments, genetics, inorganic chemistry, and mechanics.

Relationship Between the Discipline to Which Scientific Terms Belong and the Theme of the Articles. The chi-square test of association indicated that there is a significant relationship between the discipline to which the terms of prerequisite knowledge belong and the theme of the article in which they appeared, both for the magazines ($\chi^2 = 278.18$, df = 25) and the newspapers ($\chi^2 = 734.49$, df = 25). In both cases the degree of association is not very strong, .31 for the magazine terms and .34 for the newspaper terms. This implies that articles with a specific theme do not necessarily contain one type of term while articles with another theme another type of term and so on. This fact can be clearly seen from Tables 7 and 8, where for each category of articles the percentages of terms belonging to the different disciplines are expressed. It is further noted from these tables that the major disciplines of biology, chemistry, and physics serve as basic sources of terms of prerequisite knowledge for all types of articles although with different proportions. The understanding of any one type of article requires a certain level of knowledge in each of these three disciplines. On

Table 7.—Percentage of Items of Prerequisite Knowledge From Academic Disciplines for Each Theme Category of Magazine Articles

	Themes					
Science Discipline	Life	Agricultural	Mechanical	Space	Earth	Matter & Energy
Astronomy	0.4	0.0	15.2	26.5	0.0	6.7
Biology	72.7	40.4	18.2	4.7	21.8	18.0
Chemistry	14.6	40.4	21.2	21.9	45.6	30.3
Engineering	0.7	0.0	3.0	9.4	2.2	3.4
Geology	0.0	1.9	0.0	0.0	4.3	0.0
Physics	11.6	17.3	42.4	37.5	26.1	41.6
Total Number of Terms	275	52	33	64	46	89

Table 8.—Percentage of Items of Prerequisite Knowledge From Academic Disciplines for Each Theme Category of Newspaper Articles

Science Discipline	Themes					
	Life	Agricultural	Mechanical	Space	Earth	Matter & Energy
Astronomy	1.0	0.0	6.3	26.8	4.9	4.5
Biology	73.6	42.3	13.6	4.1	12.7	16.3
Chemistry	14.1	31.0	14.2	18.7	29.6	20.8
Engineering	2.2	1.4	25.6	12.6	10.5	6.2
Geology	.2	7.0	1.7	4.5	12.7	3.9
Physics	8.9	18.3	38.6	33.3	29.6	48.3
Total Number of Terms	496	71	178	198	142	178

the other hand, knowledge of astronomy terms is basic only to the reading of articles dealing with space science. Engineering terms comprise a noticeable proportion of prerequisite terms appearing in articles dealing with mechanical and earth sciences.

Relationship Between the Discipline to Which Scientific Terms Belong and the Country in Which Newspapers Are Published. No attempt was made to study the relationship between discipline and country in the case of terms appearing in magazines, because the selected magazines do not represent proportionately the countries in which they are published. On the other hand, concerning terms appearing in newspapers it was found that there is a significant difference between distributions of terms according to academic disciplines for the separate countries ($\chi^2 = 117.36$, df $= 35$). However, the degree of association between the discipline and the country attributes is very low ($\phi' = .12$). This weak association can be visually discerned from Table 9, where no particular discipline or combination of disciplines is heavily pronounced as associated with one country and not another. It is, however, noted that for Egypt, Lebanon, and Saudi Arabia, the prerequisite terms are about equally distributed between biological and physical sciences while for the rest of the countries the ratio between the physical and the biological sciences is about two

Table 9.—Percentage of Terms of Prerequisite Knowledge From Academic Disciplines for Separate Countries in Which the Selected Newspapers Are Published

Science Descipline	Egypt	Iraq	Jordan	Kuwait	Lobanon	Arabia	Sudan	Syria
Astronomy	6.9	20.3	19.8	7.9	9.4	4.7	8.3	7.4
Biology	46.8	30.0	25.9	33.3	41.8	50.5	23.6	35.5
Chemistry	16.3	15.8	12.3	14.0	17.3	20.0	12.5	19.8
Engineering	6.6	8.5	6.2	5.3	10.4	6.2	11.1	4.2
Geology	2.1	5.1	2.5	6.2	1.6	1.9	4.2	3.3
Physics	21.3	20.3	33.3	33.3	19.5	16.7	40.3	29.8
Total Number of Terms	436	177	81	114	318	210	72	121

Table 10.—Number and Percentage of Items of Prerequisite Knowledge Classified According to Linguistic Scheme

	Magazines		Newspapers	
	No.	%	No.	%
Type A	70	15.9	171	17.0
Type B	20	4.5	34	3.4
Type C	250	56.7	662	65.7
Type D	28	6.3	5	.5
Type AC	32	7.3	59	5.8
Type BC	41	9.3	77	7.6
Totals	441	100.0	1008	100.0

to one. In addition, Iraq and Jordan stand out as the only countries in which astronomy terms constitute an appreciable proportion of items of prerequisite knowledge.

Linguistic Nature of Scientific Terms. It can be noted from Table 10 that scientific terms that are genuinely Arabic (Types C and D) constitute about 65 percent of all items of prerequisite knowledge while the rest are either completely borrowed from another language (about 20%) or partially borrowed (about 15%). The majority of the loan-words are transliterated terms that did not undergo any structural modifications.

This proportion of loan-words seems to be much higher than that in the common literature. Issawi (5), as a result of the analysis of words in three Arabic novels, reported that loan-words constitute about 1.5 percent of the total number of different words. Landau (7), after studying twenty pages taken from each of sixty Egyptian books found 115 foreign words out of a total of 11,284 words (i.e., 1.0%). In another study quoted by Landau and Issawi (7) and based on 5,981 items drawn from *El-Ahram*, and *Filastin*, only fifty-one foreign words were found, i.e., 0.8 percent of the total.

The high percentage of scientific loan-words in contrast with the percentage of loan-words in the vocabulary of everyday life indicates that the resistance of the Arabic language to borrowing does not apply to scientific terms. On the contrary, scientific loan-words are incorporated into the Arabic language on the principle that it is a *must* when these terms "are widespread and are used in the scientific communities" (2:33), and that this procedure is in accordance with the way scientific Arabic developed historically. In an introduction to a compilation of scientific terms established by the Academy of Arabic Language in Cairo, it is stated that the ancient Arab scholars

> *did not care whether a scientific term was a genuinely Arabic or an Arabicized foreign one. They may even have preferred the foreign term whenever it was closer to the meaning and more perfect in expression* (1:3).

On the other hand, it is difficult to attribute the large proportion of scientific loan-words to the hastiness of journalists in translating scientific articles, assuming that it is easier to transliterate than to find a genuine Arabic term to express a new scientific concept. When the scientific terms used by the Egyptian press were compared with those published by the Academy of Arabic Language in Cairo, it was found that of the 375 items appearing in both sources, only in six cases did the press use loan-words when the Academy used genuine Arabic terms, while the press used Arabic words in the case of three terms for which the Academy used loan-words.

The Relationship Between Language and Discipline. There is a significant but weak association between the discipline and language type attributes for scientific terms appearing in magazines ($\chi^2 = 85.26$, df $= 25$, $\phi' = .19$) as well as

Table 11.—Percentage of Magazine Terms Classified According to Linguistic Type for Each Discipline

Linguistic	Disciplines					
Type	Astronomy	Biology	Chemistry	Eng'g	Geology	Physics
Type A	0.0	11.2	34.4	16.7	100.0	9.3
Type B	0.0	4.5	5.4	16.7	0.0	3.5
Type C	87.0	62.9	31.2	25.0	0.0	66.3
Type D	0.0	7.6	5.4	0.0	0.0	7.0
Type AC	8.7	6.2	10.7	8.3	0.0	5.8
Type BC	4.3	7.6	12.9	33.3	0.0	8.1

Table 12.—Percentage of Newspapers Terms Classified According to Linguistic Type for Each Discipline

Linguistic Type	Disciplines					
	Astronomy	Biology	Chemistry	Eng'g	Geology	Physics
Type A	13.0	13.5	38.8	12.9	16.2	7.9
Type B	0.0	2.1	4.4	9.4	8.1	2.8
Type C	79.7	77.6	31.1	53.0	51.4	73.1
Type D	0.0	.2	0.0	3.5	0.0	.5
Type AC	1.5	1.9	14.8	4.7	8.1	6.9
Type BC	5.8	4.7	10.9	16.5	16.2	8.8

newspapers ($\chi^2 = 209.74$, df = 25, $\phi' = .20$). In both cases, and as indicated by Tables 11 and 12, the proportion of borrowed words is lower for astronomy, biology, and physics than it is for chemistry and engineering. Further analysis shows that the low proportion of borrowed words in astronomy is due to two facts. First, many of the astronomy terms that were used belong to times when the Arabs were actively involved in the making of astronomy and consequently were used by classical Arab scholars, e.g., *nayazik* (meteors); *kawkab* (planet); *kusouf* (eclipse). Second, the terms that deal with new concepts are mainly descriptive, and therefore, could be easily expressed in Arabic, e.g., *mahattat fadaiyyat* (space station); *saroukh rafi'* (lifting rocket); *madaar* (orbit).

In the case of biology, the few loan-words used were in connection with newly discovered diseases and disorders and with biochemistry which calls for a number of scientific "proper" names. Similarly the physics loan-words expressed new ideas in electricity and atomic physics.

The high proportion of foreign chemistry words is due to the fact that names of chemicals are looked upon as universal "proper" names, and consequently are transliterated with no modifications. In the case of names of chemicals that are made of two or more words, only the descriptive part is translated, e.g., *hamid et-tannic* (tannic acid); *Khamis okseed el-fosfoor* (phosphorous pentoxide); *amlah el-calcium* (calcium salts).

The Relationship Between Language Type and Country in Which Newspapers Are Published. No significant association was found to exist between the language type attribute and the country attributes ($\chi^2 = 34.96$, df = 35). This means that the distribution of terms according to their linguistic structure does not significantly differ from one country to another. This result refutes the widespread opinion that the highly nationalistic and anti-western countries in the Arab Middle East are more zealous about the Arabic language and refuse to admit into it foreign words. It seems that the incorporation of scientific words is not a function of the political and cultural attitude of the country, as much as it is a function of the nature of the concepts that are being verbalized. (See previous section.)

Difference Between Established and Press Terminologies. To compare the scientific terminology used by the press with established terminology, Egypt was selected as an example. The scientific terms established by the Academy of Arabic Language were published in the form of annual English-Arabic lexicons entitled *Majmou'at el-Mustalahat el-'Ilmiyyat wal Fanniyyat el-Leti Aqarraha el-Majma'* (Collection of Scientific and Technical Terms Established by the Academy). The English equivalents of the scientific terms used by the press were looked up in these collections and the respective Arabic terms used in both sources were compared.

It is noted from Table 13, which is a summary of the findings, that 24 percent of the terms used in magazines and 28 percent of the terms used in newspapers had no equivalent in the Academy's *Collections.* These terms partly belong to astronomy, a discipline that is not tackled by the Academy in the existing *Collections.* The other terms used by the press with no established counterparts deal mainly with (a) diseases and treatment, (b) ecology, (c) electricity, (d) atomic and nuclear physics, and (e) engineering. Only a small number of these terms are related to recently developed concepts like electronic brain, nuclear propulsion, and sound prints while the rest express concepts that have been known for a long time.

It is further noted that about 70 percent of the terms used by the press are in agreement with the terminology established by the Academy, while only about 6 percent are different. This means that in the case of the already established terms the press did not deviate from these established terms.

These results indicate that, even in places where there are scholarly attempts to establish scientific terms in Arabic, whether by means of translating or borrowing, the journalist is still playing a significant role in the origination or establishment of scientific vocabulary. In some instances he must select the appropriate terms out of many that are in use, while in other instances he has to coin or transliterate a term that deals with a new development which he is covering in a news item. In the case of Egypt, for example, the journalist, in more than one-fourth of the cases, had either to originate or to confirm a term before the Academy of Arabic Language did so.

Consistency in the Usage of Terms. Four types of inconsistencies in scientific vocabulary were detected in the press of different Middle Eastern Arab countries, or even within one country. The first type involves the usage of a genuinely Arabic term at one time and a loan-word at another time for the same concept. Following is a set of examples of this duality in terminology:

English Term	*Loan-Word*	*Genuinely Arabic*
Acid	Aceed	Hamid
Aerial	Aerial	Hawa'i
Ammonia	Ammonia	Noshadir
Anemia	Anemia	Foqr Dem
Electrode	Electrode	Qotb
Polymer)		Ladayin
Polymerization)	Balmarat	
Microscope	Microscope	Mijhar
Vaccine	Vaccine	Loqah
Varicose	Vareez	Marad ed-Dawali

The second type of inconsistency is due to the borrowing of words from different sources. For example, *nitrogeen* is borrowed from English while *azote* is borrowed from French. The same thing applies to *hidrogeen* (hydrogen) and *idrogeen* (hydrogene, French); *bacteria* (bacteria) and *bacteer* (bacterie, French). The words for *dysentery* are *dyzantari* and *dosontaria,* which is probably derived from the original Latin word.

Inconsistencies of the third type involve the usage of more than one genuinely Arabic word for the same item. For the idea of electric or magnetic field, for example, *haql and majal* are used. For calorie three words are used, *harirat, si'ir,* and *wihdat harariyyat,* while for the aorta three different terms are used, namely, *shiryan taji, shiryan iklili,* and *'irq abhar,* in addition to a partially borrowed term, *shiryan aworti.*

The last type involves different ways of spelling the same word. Basically, this case is tied up with the use of vowels in the transliteration operations. Such differences in spelling do not lead to different words, but when the same word is spelled two or three different ways in the same article it can become confusing to the reader.

The first three types of inconsistencies in scientific vocabulary used in the Middle Eastern press put an extra burden on the reader to familiarize himself with different words that express the same meaning. Moreover, if a reader is scientifically literate to be able to read and understand the press science in one Middle Eastern Arab country he will find some difficulty understanding the press science of another Arab country. If nothing is done, this problem of multi-terminology is likely to become worse. With the proliferation of science and the development of new scientific concepts and technological inventions more coined or transliterated terms have to be

Table 13.—Comparison of Scientific Terms Used by the Egyptian Press With Terms Established by the Academy of Arabic Language in Cairo

Science	Magazines				Newspapers			
Discipline	Alike	Different	Press only	Total	Alike	Different	Press only	Total
Astronomy	1	0	12	13	2	1	27	30
Biology	23	2	2	27	153	13	38	204
Chemistry	16	1	0	17	63	2	6	71
Engineering					9		20	29
Geology					5		5	10
Physics	15	2	5	22	60	7	26	93
Total	55	5	19	79	292	23	122	437
Percentage	69.6	6.3	24.1	100.0	66.8	5.3	27.9	100.0

introduced into the Arabic scientific literature. Unless there are some guidelines or an overall plan for the selection of these terms there will be a chaos of terminology, which will be detrimental to the communication and activity of science in the Arab world.

Conclusions and Implications

It is concluded from this study that the reading of scientifically and/or technologically oriented articles in the Arab Middle Eastern press requires a considerable amount of prerequisite knowledge in the various disciplines of science. Articles belonging to any theme type require knowledge in each of the three major disciplines of science, namely, physics, chemistry, and biology. Since formal instruction in schools is the major source of these items of prerequisite knowledge, it follows that different or combined courses in biology, chemistry, and physics must be introduced into the school curriculum at about the same time if we want such courses to be instrumental in making the individual scientifically literate. Such course or courses must also include topics in ecology, organic chemistry, and atomic and nuclear physics.

It is further concluded that a high proportion of the scientific vocabulary of the press is either totally or partially borrowed. The distribution of terms according to their linguistic structure depends upon the discipline to which they belong and not upon the country in which they are used. The process of borrowing led to the introduction of different technical terms to convey the same meaning, depending on the language from which the term was borrowed. The same problem exists when more than one translation is given for the same term.

The situation might have been easier had the usage of scientific terms by the press been preceded by a scholarly study and approval by one or more academies of Arabic language. In that case, a certain plan could have been devised to guide the work of such academies. But since a considerable proportion of the press terminology has to be decided upon by the journalists they should agree on some guidelines that govern their procedures in borrowing or translating scientific terms. Such guidelines should include the following principles:

1. Transliteration must be employed when the term dealt with is universally used. In this case the original form and not the modified forms of modern languages should be transliterated.
2. Terms that differ with different languages (e.g., rating, aerial), or terms that do not carry any special scientific meaning in their form (e.g., filter), should not be transliterated but translated. The operation of translation should be according to the scientific and not the apparent meaning of the term. Consequently *zar'* and *irsal* are poor translations of *transplantation* and *transmission*.

In addition, academic institutions must be more active in originating and establishing appropriate scientific terms by anticipating the scientific developments that are apt to make headlines in the press. On the other hand, more scholarly research is needed in the areas of compilation, comparison, categorization, and statistical analysis of existing terms used in the different Arab Middle Eastern countries so that the scientific vocabulary in Arabic can serve its purpose more fully.

FOOTNOTES

1. This article is part of a study entitled, *Scientific Literacy in the Middle East,* and supported by a grant from the Middle East Area Program, American University of Beirut, with funds provided by the Advanced Research Projects Agency.
2. Science content analysis of newspapers and magazines is the subject of another study presently conducted by the author.

REFERENCES

1. Academy of Arabic Language, *Majmou'at el-Mustalahat el'Illmiyyat wal Fanniyyat el-Leti Aqarraha el-Majma'* (Arabic), *Collection of Scientific and Technical Terms Established by the Academy,* Volume 2, Cairo, 1960, pp. 3, 4, 6 (translated).
2. Fahmi, Hasan Husein, *El-Marja' fi Ta'reed el-Mustalahat el- Ilmiyyat wal Fanniyyat wal Handasiyyat* (Arabic), *A Reference Book for the Translation of Scientific, Technical, and Engineering Terms,* En-Nahda el-Masriyya Press, 1958, p. 33 (translated).
3. Hays, William L., *Statistics for Psychologists,* Holt, Rinehart, and Winston, New York, 1963.
4. Hines, Neal O., "Atomic Energy and the Press," *Journalism Quarterly,* 24:315–322, 1947.
5. Issawi, Charles, "European Loan-Words in Contemporary Arabic Writing: A Case Study in Modernization," *Middle Eastern Studies,* 3:110–133, January 1967.
6. Koelsche, Charles L.; Morgan, Ashley G., Jr., *Scientific Literacy in the Sixties,* University of Georgia, Athens, 1964.
7. Landau, Jacob, A Word Count of Modern Arabic Prose, quoted in Issawi, Charles, "European Loan-Words in Contemporary Arabic Writing: A Case Study in Modernization," *Middle Eastern Studies,* 3:124, January 1967.
8. Pella, Milton O.; O'Hearn, George T.; Stiles, Lindley J., *Scientific Literacy,* University of Wisconsin, Madison, 1966.
9. Reeves, Richard, "A Survey of Science Reporting in Representative American Newspapers," unpublished master's thesis, Boston University, 1950.
10. Rowe, Edward, "A Study of Science News and Its Readership in Representative American Newspapers," unpublished master's thesis, Stanford University, Stanford, California, 1951.
11. Withey, Stephan B., "Public Opinion About Science and Scientists," in Barber, Bernard; Hirsch, Walter (eds.), *The Sociology of Science,* The Free Press of Glencoe, New York, 1963, p. 155.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

Academic Performances of Matched Groups of Community College and 4-Year College Students

LINCOLN H. HALL

ABSTRACT

A total of 180 June, 1964, high school graduates, ninety of whom entered a public junior college and ninety of whom entered 4-year institutions directly from high school, were studied through June, 1969. The two groups were matched in terms of sex, IQ, aptitude test scores, high school grade point averages (GPA's), and high school majors. Although the junior college students attained higher mean GPA's in the first 2 years of college (2.770 as compared with 2.368 on a 4.00 scale), a significant difference was not found in the upper division years, the native students earning a mean GPA of 2.747 and the junior college transfers a mean of 2.742. Of the initial groups of ninety each, fifty-five, or 61.2 percent, of the 4-year college freshmen received bachelor's degrees as compared with sixty-one, or 67.1 percent, of the junior college freshmen after 5 years.

A CONTINUING concern of community college educators is the academic performance of their former students at 4-year institutions of higher learning. One of the more commonly employed techniques for judging the effectiveness of the community colleges' transfer programs is to compare the relative scholastic performances of transfer students in their 2-year and 4-year institutions. Knoell and Medsker (1) found that between 75 and 80 percent of the junior college transfer students achieved their degree objectives during a 4-year period following their transfer to a 4-year college or university in 1960. Thornton (3) reports a study by Stickler which found that the total group of all Florida community college transfers combined approached the same grade point averages (GPA's) in the university that they had earned in the community college before transfer, with the mean GPA being only 0.11 grade points lower in the university.

Community college transfer programs are also evaluated by comparing the relative academic performances of transfer and native students at 4-year institutions. Medsker (2) cited studies which indicated that, in terms of both persistence and college grades, native students at 4-year colleges performed only slightly better than did community college students in the transfer institutions.

In none of the studies cited above or reviewed in the literature was an attempt made to equate the students who were being compared in terms of either intelligence, college aptitude, or high school GPA.

Purpose and Method of the Study

The purpose of the present study was to compare the academic performances and persistence of two matched groups of college-bound June, 1964, high school graduates, one comprising students who entered a California public community college and the other consisting of students whose college careers were begun in 4-year colleges and universities.

Selection of Ss was designed to equate the two groups in terms of sex, intelligence, scholastic aptitude, type of program taken in high school, and high school GPA's. Inclusion of high school GPA's was assumed to provide some indication of students' motivation when combined with the measures of intelligence and academic aptitude.

During the spring of 1969, principals in the twelve high schools which lie within the service area of the College of the Sequoias, a community college in California's central valley, were asked to provide the names, high school transcripts, and colleges of attendance of all 1964 graduates who had entered 4-year colleges and universities following their high school graduation.

All high schools in the community college district participate in a cooperative testing program under which a group intelligence test, the Lorge–Thorndike, and a scholastic aptitude test, the

Differential Aptitude Test Battery (DAT), are administered in the tenth and eleventh grades. The test results are provided to the community college where they are transferred to profile sheets, copies of which are returned to the high schools for use by counselors. As a consequence, intelligence and scholastic aptitude test data for virtually all high school graduates in the college district were available for use in the present study.

Each of the 4-year college and university students was matched with a community college incoming freshman on the following bases:

1. *Sex*
2. *Intelligence quotient. Subjects were matched within 5 IQ points.*
3. *Combined verbal and numerical percentile ranking on the DAT battery. A maximum variance of 5 in percentile rankings was allowed.*
4. *High school major. All subjects had pursued college preparatory programs in high school.*
5. *High school GPA. Subjects were matched within 0.2 of a grade point on a 4.0 scale.*

In addition to all students in the two groups having followed college preparatory programs while in high school, the community college students had entered transfer programs in the 1964 fall semester.

A total of ninety-five names was provided by the high schools. Each of these students was matched with an entering community college freshman in accordance with the five criteria previously cited. In the summer of 1969, letters describing the study were written to officials in the 4-year institutions which the 1964 high school graduates had entered and those to which the community college students had transferred requesting transcripts of the Ss' scholastic records. In addition, the institutions were asked to provide the names of colleges to which the students had transferred in order that transcripts of work at these institutions might be obtained. Four-year institutions attended by five of the students, four who had entered from high school and one who was a community college transfer, refused to release the students' records. Accordingly, in each case, the matched pair was removed from the study, resulting in a final total of ninety students in each group.

GPA's were calculated for students during their attendance at the community college and for the first 2 years, or portion thereof, of the 4-year college students' attendance. In addition, GPA's were computed for the semesters following the first 2 years until graduation or termination of college enrollment for all students.

Findings

Table 1 presents mean intelligence quotients, combined verbal and numerical aptitude test percentiles, and high school GPA's for the two groups, each of which contained forty-seven men and forty-three women. Testing the null hypothesis, there were no significant differences between the groups in the three selection criteria.

Academic performances are shown in Table 2. A significant difference favoring the community college students by .402 grade points, 2.770 as compared with 2.368, was found between the means of the two groups during the first 2 years. An insignificant difference favoring the native students by .005 grade points, 2.747 as compared with 2.742, was found for the upper division years.

Table 1.—Mean IQ's, Aptitude Test Results, and High School GPA's of 4-Year and Community College Freshmen, Fall 1964

	4-year Institutions	Community Colleges	"t" values
IQ	111.036	111.283	0.1243
Verbal and Numerical Percentiles	74.250	75.262	0.420
High School GPA's	3.304	3.286	0.3000

The mean GPA's of the community college students who transferred to 4-year institutions declined from 2.804 to 2.742, between community college and 4-year institution. Comparing first semester or quarter grades at their transfer institutions with those at the community college, a "transfer shock" of 0.22 grade points, a decline from 2.804 to 2.584 in the first term, was found.

Table 2 also shows that of the ninety students who entered 4-year institutions, only sixty-six, or 73.3 percent, continued into the junior year as compared with seventy-six or 84.4 percent, of the community college students. Baccalaureate degrees were earned by fifty-five, or 61.1 percent of the original 4-year students and sixty-one, or 67.8 percent, of the students who entered a community college. The fifty-five native students and sixty-one community college students who received bachelor's degrees represented 83.3 percent and 80.3 percent, respectively, of the Ss who entered the junior year at a 4-year institution.

Table 2.—Lower Division and Upper Division Academic Performances of 4-Year and Community College 1964 Entering Freshmen: 1964–1969

	Lower Division GPA	Upper Division GPA	Bachelor's Degrees Earned	
			Number	Percent
4-year College Freshman	2.368 (N=90) *t= 4.96	2.747 (N=66) t=0.0571	55	61.18
Community College Freshman	2.770 (N=90)	2.742 (N=76)	61	67.06

$*p < .01$

In Table 3, upper division academic performances of the sixty-six native and seventy-six transfer students are compared by type of 4-year institution: California State college, University of California campus, and private college. Slightly higher GPA's were earned by the transfer students at University of California campuses and private colleges while native students earned higher GPA's at the state colleges. In neither instance were the differences statistically significant. Proportionately more native students who continued into the upper division level earned baccalaureate degrees than did transfer students at state colleges and private colleges whereas a higher percentage of transfer students earned degrees at University of California campuses.

Table 3.—Upper Division Academic Performances of Native and Transfer Students by Type of Institution

Type of Institution	Student Classification	Upper Division GPA	Bachelor's Degrees Earned	
			Number	Percent
California State College	Native (N=32)	2.830 (t= 0.550)	27	84.37
	Transfer (N=48)	2.750	38	79.17
University of California	Native (N=15)	2.556 (t= 0.107)	12	80.00
	Transfer (N=17)	2.664	14	87.50
Private College	Native (N=19)	2.754 (t= 0.098)	16	84.21
	Transfer (N=11)	2.881	9	81.82

Conclusions and Discussion

Community college students of comparable intelligence, academic aptitude, and high school GPA performed as well academically in their upper division work as did students who began their college careers at 4-year institutions.

Persistence of community college students in transfer programs during the freshman and sophomore years was greater than that of students who entered 4-year colleges following high school graduation. The proportion of community college students who attained their degree objectives exceeded that of native students by almost 6 percent. However, a slightly higher proportion of native students persisted from the beginning of the junior year to completion of bachelors degree requirements. The evidence suggests strongly that some students whose educations were terminated during the first 2 years of attendance at 4-year colleges would have atttained their bachelor's degree objectives if they had begun their college educations at a community college.

Although "transfer shock" was experienced in the first semester by the community college students, their academic performances in succeeding semesters improved sufficiently to equal that of native students in their upper division work.

Both groups of students attended predominantly public institutions with sixty-five, or 85.5 percent, of the community college students transferring to California State colleges and University of California campuses and sixty-eight, or 75.6 percent, of the original ninety four-year college students entering California public institutions.

REFERENCES

1. Knoell, Dorothy M.; Medsker, Leland L., *From Junior to Senior College: A National Study of the Transfer Student*, American Council on Education, Washington, D. C., 1965, pp. 26–28.

2. Medsker, Leland L., *The Junior College: Progress and Prospects*, McGraw-Hill Book Company, Inc., New York, 1960, pp. 323–325.

3. Thornton, James W., Jr., *The Community Junior College*, John Wiley and Sons, Inc., New York, 1960, pp. 64–65.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 64, Number 10, July–August 1971)

The Effect of Selected Media Feedback Upon the Interactive Behavior of Student Teachers[1]

B. CHARLES LEONARD, FREDERICK JOHN GIES, and **JON S. PADEN**
University of Missouri—Columbia **I/D/E/A/**

ABSTRACT

The problem involved in this study was to ascertain whether student teachers receiving supervisory feedback with the aid of video and/or audio tape replay demonstrate a greater change in their classroom verbal interactive behavior than student teachers who receive supervisory feedback with no video or audio feedback. Evidence indicates that those receiving video feedback utilize less "direct teaching influence" and more "indirect teaching influence" as categorized by Flanders' Verbal Interaction Analysis System. Within the audio group, the magnitude of change in the student teachers' instructional behavior was not statistically significant. The magnitude of the directional change was greater for those students receiving video or audio recorded feedback than the magnitude of change in the control group.

THE PURPOSE of this investigation was to analyze three distinct supervisory feedback techniques in relation to their effect on the interactive behavior of student teachers. The specific research questions postulated were:

1. *Will the act of seeing and hearing oneself teach via video tape recordings effect a greater change in one's verbal teaching behavior than simply discussing the teaching activity with a supervisor?*
2. *Will the act of hearing oneself teach via audio tape recordings effect a greater change in one's verbal teaching behavior than simply discussing the teaching activity with a supervisor?*
3. *Will the act of seeing and hearing oneself teach via video tape recordings effect a greater change in one's verbal teaching behavior than hearing oneself teach via audio tape recordings while discussing the teaching activity with a supervisor?*

Methodology

Standard research procedures were utilized in conducting this study. In both the fall and winter semester twelve student teachers of English were selected, matched on grade point average (GPA), and assigned to one of three groups: (1) video experimental; (2) audio experimental; or (3) control. During the semester, each student teacher was audio and video taped in the process of teaching. At a minimum, such data were collected before and after the treatment procedure. During the course of both semesters, each student teacher received feedback from her respective supervising teacher, according to her assigned group. The experimental variable for the study was the measured effect the three different kinds of feedback had upon the classroom verbal interaction of the student teachers. The verbal interaction was measured and codified by use of Flanders' Verbal Interaction Analysis System (1). Two supervisors of English classes at the Laboratory School, University of Missouri–Columbia, were utilized in this study. Each supervisor worked with six student teachers each semester of the school year 1968–69.

The research design included collection of pretreatment and post-treatment data for both experimental groups and the control group. The crux of this investigation rested on the measured amount of the demonstrated change of classroom verbal behavior for each group. Since the questions which this study attempted to answer dealt with the direction and magnitude of change, all data were subjected to the two-tail test of significance at the .05 level of confidence.

The problem was to ascertain the medium, i.e., video, audio, or regular supervisory feedback, that had the greatest impact on changing the classroom interactive behavior of the student teachers as measured by Flanders' Interaction Analysis Categories. In the research hypotheses tested, each assumed no difference in student teachers' interactive behavior change, by Flanders' categories, from pretest to posttest data associated with video, audio, or control groups.

Since the concern in this investigation was for the relative magnitude as well as the direction of the differences, the Wilcoxen matched-pairs signed-ranks test seemed to be the most useful tool (3:75). Utilizing this method, the null hypothesis assumes the sum of the positive ranks will equal the sum of the negative ranks. If, on the other hand, a difference of magnitude is found between sums, it constitutes evidence for the rejection of the null hypothesis that the two sets of measurements are from the same population (2:360).

Comparison of the three groups was aided by the computation of the Kruskal-Wallis one-way analysis of variance (3:184–94). This application is consistent with the use of nonparametric statistics with more than two samples. It has power-efficiency of $\frac{3}{\pi} = 95.5$ percent, when compared with the F-test, the most powerful parametric test (3:194).

Findings

Table 1 summarizes the findings resulting from the application of the Wilcoxen test to the combined data from Flanders' Categories 1, 2, 3, and 4, which constitute the index of "indirective teaching influence."

In analyzing the data in Table 1, the most pertinent finding was the significant increase in frequency of "indirect teaching influence" demonstrated by the video experimental group. The change, in this case gain, was found to be significant at the .01 level of confidence. The measured gain scores for this variable were not statistically significant for the audio experimental and control groups.

Combining the data from Categories 8 and 9 produced the index of "student talk." As indicated in Table 2, the change and/or gain, was in a positive direction for all three groups. However, the amount of gain proved to be statistically significant in only the video experimental group. The calculated Wilcoxen T value of 4 for the video group was significant at .05 level of confidence.

Table 1.—Comparison of Wilcoxen T Values for Increase in Non-directive Teaching Influence as Coded by Flanders' Verbal Interaction Categories 1, 2, 3, and 4 Between the Groups

Groups	Number of Student Teachers	Positive Value	Negative Value	Calculated T Value[a]
Video	8	35	1	1*
Audio	8	16	20	16
Control	8	29	7	7

[a] T Value required for significance at .05 level was 4 or less for each group.

*T Value statistically significant at the .05 level of confidence.

Table 2.—Comparison of the Wilcoxen T Values for Increase in Student Talk as Coded by Flanders' Verbal Interaction Categories 8 and 9 Between the Groups

Groups	Number of Student Teachers	Positive Value	Negative Value	Calculated T Value[a]
Video	8	32	4	4
Audio	8	31	5	5
Control	8	31	5	5

[a] T Value required for significance at .05 level was 4 or less for each group.

"Direct teacher influence" which consists of data from combined categories 5, 6, and 7 decreased during the investigation period for all three groups. While a decrease in "direct teacher influence" resulted for all groups, only the decrease found within the video experimental group was significant at the .05 level of confidence.

Table 3.—Comparison of Wilcoxen T Values of Decrease in "Direct Teacher Influence" as Coded by Flanders' Verbal Interaction Categories 5, 6, and 7 Between the Groups

Groups	Number of Student Teachers	Positive Value	Negative Value	Calculated T Value[a]
Video	8	2	34	2
Audio	8	10	26	10
Control	8	7	29	7

[a] T Value required for significance at .05 level was 4 or less for each group.

Since the Flanders' Verbal Interaction instrument classifies the total verbal interaction taking place in a classroom into one of ten descriptive categories, it seemed logical to assume that if "student talk" and "indirect teacher influence" increased significantly for the video experimental group, then "teacher lecture" would decrease significantly. Table 4 presents data to substantiate this assumption. The decrease in "teacher lecture," Flanders' category 5 for the video group was found to be significant at the .05 level. Comparisons of the pretreatment scores with the posttreatment scores for the audio experimental and control groups were not significant.

Table 4.—Comparison of Wilcoxen T Values for Decrease in Teacher Lecture as Coded by Flanders' Verbal Interaction Category 5 Between the Groups

Groups	Number of Student Teachers	Positive Value	Negative Value	Calculated T Value[a]
Video	8	2	34	2
Audio	8	10	26	10
Control	8	9	27	9

[a] T Value required for significance at .05 level was 4 or less for each group.

Conclusions

Based upon the previously stated findings, the following conclusions appear justified.

1. *The technique of video recording student teachers in the act of teaching, and utilizing the recorded teaching sequence as feedback as a supplement to the regular supervisory procedures, tends to change some aspects of the student teachers' instructional behavior. Specifically, the evidence indicates that the student teachers receiving video feedback utilize less "direct teaching influence" and more "indirect teaching influence" as categorized by Flanders' Verbal Interaction Analysis System.*
2. *Utilizing the process of audio recording student teachers in the act of teaching and subsequently studying the recorded teaching sequence with the supervising teacher, effects some change in the instructional behavior of the student teacher. However, it was evidenced in this study, that within the audio group, the magnitude of change in the student teachers' instructional behavior was not statistically significant.*
3. *The magnitude of the directional change is greater for those student teachers receiving video or audio recorded feedback than the magnitude of change in the control group. However, the difference in the magnitude of change in each of Flanders' ten categories between the three groups was not statistically significant.*

Acknowledging the need to provide more effective supervisory feedback to aid student teachers in their professional training, it seems reasonable to infer from this study that utilizing video recorded classroom interaction as part of the supervisory process produces positive results.

FOOTNOTE

1. The research reported herein was supported through the Cooperative Research Program, Small Grants of the Office of Education, U. S. Department of Health, Education, and Welfare.

REFERENCES

1. Amidon, Edmund J.; Flanders, Ned, *The Role of the Teacher in the Classroom,* Association for Productive Teaching, Minneapolis, Minnesota, 1967.
2. Ferguson, George A., *Statistical Analysis in Psychology and Education,* McGraw–Hill, New York, 1966.
3. Siegel, Sidney, *Nonparametric Statistics for the Behavioral Sciences,* McGraw–Hill, New York, 1956.

RESEARCH BRIEFS—Continued from page 442

districts they exceeded the state average by $137.00 and $20.00. Taken together school districts with humanities programs spent some $54.00 more per pupil than the state average.

The humanities are very broad. Areas of knowledge included in over one-third of the programs were literature, social studies, art, and music. Areas included in over 10 percent were drama, religion, and architecture. Other areas mentioned by smaller percentages were dance, science, and Latin. Various literary sources used ranged from Aeschylus and Plato to John Kerouac and Marshall McLuhan. Some eighty different works were listed as required reading.

Teaching techniques and procedures used in humanities programs tend to emphasize individual effort, and some degree of intellectual freedom for the student. Procedures used in over 50 percent of existing programs are independent study, oral reports, media center work, small group work, student teacher planning, team teaching, term papers, and Socratic questioning.

In the state of Ohio, the following generalizations can be made regarding humanities programs:

(1) *These are found in all kinds and sizes of school districts;*

(2) *Most programs in the humanities are secondary, but a few are elementary;*

(3) *School districts with programs in the humanities tend to spend more per pupil than other districts of the same type; and*

(4) *Programs in the humanities tend to emphasize a wide variety of materials and procedures, with emphasis placed upon independent study and individual effort.*—**Daniel S. Parkinson, Miami University, Oxford, Ohio.**

Footnote

The complete research report may be obtained by writing to Dr. Daniel S. Parkinson, 305 McGuffey Hall, Miami University, Oxford, Ohio 45056.

FIELD NEWS

Willard E. North
Central Missouri State College

A PROJECT INTENDED TO STIMULATE the teaching of behavioral sciences in grades 1 and 2, with the expectation that eventually behavioral sciences will be taught throughout the elementary grades, is under way at Teachers College, Columbia University. The project is titled "Teaching Children About Human Behavior" and is under the direction of Sheldon R. Roen, Associate Professor of Psychology and Education. The project is funded by the National Institute of Mental Health.

Professor Roen sees several compelling reasons for incorporating the behavioral sciences into this school program at an early age. First of all, the behavioral sciences bridges the much discussed chasm between the "two cultures" showing the concerns of both the humanities and the pure sciences. Secondly, at least as many basic, useful skills and knowledge are obtained from studying the behavioral sciences as from the humanities in pure science. The language of the behavioral sciences helps the children learn to express verbally their experiences and emotions, helps them to realize that such concepts as hostility, poverty, or power, for example, are appropriate for discussion. With the present demand for relevance in school curriculum, the behavioral sciences could provide immediate, significant relevance.

Thirdly, the behavioral science program in school can provide a major avenue of primary prevention of psychiatric disorders and maladjustment. The more children know about themselves and their total environment, according to Roen, the more able they will be to handle their lives. By legitimizing for children such topics as the nature of prejudice, how personality relates to social class and culture, defensive strategies employed by people, and disadvantaged situations, he hopes to increase their understanding and decrease their sense of alienation.

Roen and his aids, Stephen V. Gullo, a doctoral student in developmental psychology, and Claus Frank, a doctoral student in clinical psychology, plan to develop a "naturalistic curriculum" based on their study of the developmental sequence of children's interest in behavior. They will undertake a Piaget-type analysis of children's concepts in the area of understanding and behavior, prior to creating actual lessons based on their research. The creation of these lessons will be the focus of a seminar which Roen is teaching this year at Teachers College. Open only to classroom teachers of the first and second grades, the seminar will provide them, during the first semester, with the opportunity to develop lessons in the behavioral sciences, using the research findings of Roen and his staff. During the second semester, they will try out the new lessons in the classroom.

* * * * * *

TWO COMMUNITY COLLEGES HAVE JOINED the consortium working cooperatively with the Regional Educational Laboratory for the Carolinas and Virginia (RELCV) to improve classroom instruction. The expansion increases the number of 2-year institutions in the group from seventeen to nineteen. The two new colleges are the College of the Albemarle, Elizabeth City, N. C.; and Wayne Community College, Goldsboro, N. C.

RELCV conducts workshops for faculty members at the nineteen consortium schools on the "system's approach" to individualize instruction. This approach uses behavioral objectives, criterion-referenced tests, and a variety of teaching techniques to enable students to learn at their own pace. It also incorporates a process for revising instruction when students do not accomplish course objectives. The laboratory also holds conferences for trustees, presidents, deans, and other key administrators on how they can stimulate and support efforts by faculty members to individualize instruction.

Each of the cooperating colleges has appointed an Educational Development Officer (EDO) who serves as the on-campus contact for the laboratory. The EDO assists faculty members in evaluating the effectiveness of the classroom instruction and by conducting on-campus workshops on the systems approach for new faculty members. He assists administrators by conducting institutional research related to the academic program, so that administrators' decisions dealing with instruction can be based on accurate information. RELCV's Junior and Community College Division has now trained more than 1,500 faculty members in the "system's approach" to instruction. Teacher training materials developed by the laboratory have been published by Westinghouse Learning Corporation and the Educational Resources Information Center.

* * * * * *

INFORMATION SYSTEMS: CURRENT DEVELOPments and Future Expansion, the proceedings of a 1-day seminar for congressional members and their staffs on the promises and problems of information systems, is available from the American Federation of Information Processing Press. The volume covers a wide range of topics related to information systems and includes such areas as: *Computer Applications in Political Science; Communications and Future Information Systems;* and *Understanding Information Systems.* The volume contains the complete text of six papers presented at a 1-day conference in Washington, D. C., sponsored by AFIPS and the Association for Computing Machinery (ACM). The purpose of the seminar was to acquaint congressional members and their staffs with the current problems and future developments in information systems and to lay the groundwork for a continuing dialogue between the Congress and the computing and information systems profession. Contributing autors include: J. D. Aron, Federation Systems Center, IBM; Robert Lee Chartrand, Legislative Reference Library, Library of Congress; Lewis Feldner, Business Planning, On-Line Computer Corporation; Kenneth Janda, Department of Political Science, Northwestern University; Marvin Kornbluh, Lecturer and Consultant; and Andrew J. Lipinski, Research Fellow, the Institute for the Future.

Copies may be ordered from the AFIPS Press, 210 Summit Avenue, Montvale, New Jersey 07645. The cost is $5.00 per copy.

* * * * * *

LEHIGH UNIVERSITY HAS BEEN SELECTED to receive grants from the Carnegie Foundation to support development of a Doctor of Arts (DA) degree program. This program will be initiated at Lehigh beginning with the fall semester of the 1971–72 academic year. It is anticipated that the first students will receive the degree in June 1974.

The Department of Psychology, of which Dr. Francis J. Wuest is chairman, will be the first academic department at Lehigh to install the program. The Department of Government and the College of Business and Economics are currently preparing proposals for DA programs.

The Carnegie Grant, according to Dr. Wuest who has been appointed coordinator of the DA program at Lehigh, will be used for curriculum development, establishment of an intern program, and liaison with area colleges, training in interpersonal awareness, and conferences on teacher training. The DA program will focus on the preparation for teaching but will not exclude a research emphasis.

* * * * * *

SOCIAL AND TECHNOLOGICAL CHANGE: Implications for Education, edited by Phillip K. Piele, Terry L. Eidell, and Stewart C. Smith, is now available from the Center for Advanced Study of Educational Administration at the University of Oregon. This monograph is divided into five parts, each part having a different author. Each part presents a "state of the knowledge" paper devoted to some aspect of social and technological change having direct implications for education.

In Part I, Willis W. Harmon presents a broad overview of our changing society and its implications for the future of education. Richard C. Williams, in Part II, presents an analysis of internal and external conditions contributing to the rise of teacher militancy in the schools. This analysis is followed by description of three alternative models for improving the involvement of teachers in the decision-making process. In Part III, beginning the technological section of the monograph, Roger A. Kaufman seeks to promote a common understanding among administrators as to the use of the system approach in solving educational problems. Marvin C. Alkin and James E. Bruno, in Part IV, discuss applications of systems approach to educational planning. In Part V, John Evans discusses educational management information systems.

Requests for *Social and Technological Change: Implications for Education* should be addressed to Editor, ERIC Clearing House on Educational Management, University of Oregon, Eugene, Oregon 97403.

* * * * * *

A NEW PROGRAM OFFERED JOINTLY by the Teacher Corps and the Peace Corps makes it possible to obtain a teacher certificate, a master's degree, and a job overseas. The Teacher Corps and the Peace Corps have recently agreed to establish a program which will prepare teachers for programs of educational change both in this country and in developing nations overseas. In the new program, college graduates will spend approximately 1 year as Teacher Corps interns in a low-income area school while pursuing graduate study at a participating university. Following this training, 2 years will be spent as a Peace Corps volunteer overseas teaching in Asia, Africa, Malaysia, or the Philippines.

The Midwestern Teacher Corps Recruitment Center is especially interested in recruiting mathematics and science (physics and chemistry) majors to participate in the program assigned jointly to Texas Southern University in Houston and in Sierra Leone, West Africa. Following an internship in Houston, the volunteer will receive 3 months of Peace Corps orientation in Sierra Leone. Requirements for completion of the master's degree will be completed with special courses, summer seminars, and independent study while serving as a Peace Corps volunteer.

The combined 3-year service is expected to produce experienced and innovative specialists and teachers for the school systems cooperating in the internship program.

Those interested in more information and an application form should write to: Teacher Corps Recruitment and Referral, University of Nebraska at Omaha, Center for Urban Education, 3805 North 16th Street, Omaha, Nebraska 68110.

* * * * * *

FOR SOME TIME, THE TENNESSEE Higher Education Commission has been working on the development of better information for management of colleges and universities. Major efforts to date have been to get more planning information about students, and the work has been conducted within the framework of the National Management Information System being developed by the Federal Government. To guide this effort, a new committee called the "Management Information Systems (MIS) Steering Committee" has been named by the Tennessee Higher Education Commission.

The MIS Steering Committee will have the following functions: (1) to remain abreast of developments and demands at state, regional, and national levels for MIS; (2) to inform the administration of each institution of these developments; (3) to create, as needed, special committees to concentrate on each of the five management information components defined by the Western Interstate Commission for Higher Education (WICHE): students, faculty and staff, facilities, finance and courses, and a committee for investigating data processing resources and needs; (4) to provide advisory assistance to institutions and the Commission in the implementation of management information techniques. A group on Student Information Systems has already been working, and plans in this area are fairly well advanced for a statewide computerized system. Later, other task groups will work on information systems for finance, facilities, faculty and staff, and resource requirements prediction models.

Dr. Jerry H. Rust, Jr., Associate Director for Fiscal Affairs, Tennessee Higher Education Commission and Mary Mayhew, Program Analyst for the Commission are coordinating this work.

* * * * * *

A REPORT ANALYZING PROGRAMS which utilized mobile facilities to provide preschool instruction has been prepared by Southeastern Educational Laboratory. SEL *School Practices Report Number III, Preschool Instruction in Both Facilities; Description and Analysis* by Jennifer Howse should be of interest to school officials seeking alternative delivery systems for preschool and kindergarten programs.

The major portion of the report is devoted to three programs in the Southeast: SEL Redimobile Project, Appalachia Early Childhood Education Program, and the Florida Migrant Education Program. These were the only programs using mobile units to deliver instructional programs to preschool children reported in response to a survey of all states and trust territories. However, request to the preschool and elementary consultants of State Departments of Education located a number of other programs using mobile facilities for various instructional purposes. These are summarized.

Copies of the report may be obtained by writing Southeastern Educational Laboratories, 3450 International Boulevard, Atlanta, Georgia 30354.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

JOSEPH R. ELLIS is Professor of Education and Coordinator of Research and Services for the College of Education at Northern Illinois University. He received the EdD from the University of Texas. His professional experience includes teaching at Oak Ridge High School, Oak Ridge, Tennessee; Appalachian State University; the University of Texas, and Indiana State University. Dr. Ellis was the Director of this U.S.O.E. funded study. JOAN L. PETERSON holds degrees from Rockford College, Northwestern University, and the University of Nebraska. She has had public school teaching experience at the junior high school level, where she first became interested in the possibility of separate education for boys and girls. At the present time she is on the staff of Secondary Education at Northern Illinois University teaching a combined undergraduate educational psychology–adolescent psychology course. Dr. Peterson was Associate Director of the study.

NEAL R. GAMSKY is the Director of The Child Development Center at the Cooperative Educational Service Agency #13 in Waupun, Wisconsin. He received his BS in Sociology and his MS and PhD in Counseling Psychology from The University of Wisconsin. He has served as an industrial personnel executive, counselor, university lecturer, mental health consultant, director of rehabilitation services, director of pupil personnel services, and director of research. Fifteen of his articles have been published in educational journals. FAYE WILLIAMS LLOYD received her BA from Marian College and a MS in counseling and guidance and a fellowship in research and statistics from The University of Wisconsin-Milwaukee. She is currently a doctoral candidate at The University of Wisconsin and a research specialist at the Cooperative Educational Service Agency #13.

ALTON R. GODDARD is Assistant Professor of Mathematics at East Texas State University, Commerce. He received his BA from Howard Payne College, MS from North Texas State University, and PhD from Texas A & M University. In addition to having taught at the high school level in Texas, he has taught at Tarleton State College and Texas A & M University. He is presently teaching Numerical Analysis and Computer Science.

HULDA GROBMAN is Professor of Education at New York University. She holds a BA from the University of Pennsylvania, a MPA from the University of Michigan, and an EdD from the University of Florida. She has been concerned with curriculum development and evaluation for the past decade, as a staff member of the Biological Sciences Curriculum Study and the Joint Council on Economic Education and as Evaluation Consultant to the Interaction Science Curriculum Project. Her publications in recent years include books on curriculum development and evaluation. She has also co-authored achievement tests in science.

WADI' D. HADDAD is Director of the Science and Mathematics Education Center at the American University of Beirut, Beirut, Lebanon. He received his PhD from The University of Wisconsin. In addition to teaching and research activities, he has served as a science education consultant to the Ford Foundation and to a number of countries in the Middle East. He is presently directing a project of science and mathematics curriculum development and textbook writing for the junior high school level in Lebanon.

LINCOLN H. HALL is Dean of Instruction at the College of the Sequoias in Visalia, California. He received his BS from UCLA in Economics and his MA in Political Economy from the Claremont Graduate School. His PhD was granted by the University of Southern California in Higher Education. After serving for 2 years as a high school social studies teacher, Dr. Hall became an Instructor in Economics at the College of the Sequoias where he was subsequently a Counselor, Dean of Evening College, and Dean of Student Personnel before assuming his present position.

B. CHARLES LEONARD received the BS in Education from Southwest Missouri State College and the MEd and EdD degrees from the University of Missouri—Columbia. He has been a teacher and superintendent in Missouri public schools. He has served as the director of Peace Corps training at the University of Missouri and administrator-supervisor of a variety of research and development projects. Presently, he is the director of the Center for Educational Improvement, College of Education, University of Missouri—Columbia. FREDERICK JOHN GIES received a BA from De Paul University and the MEd and EdD degrees from the University of Missouri—Columbia. He has served as a junior-senior high school principal and teacher of English and Latin in Illinois and Missouri public schools. He has also served as the coordinator of Supervised Teaching and Assistant Professor of Education and Psychology at Southwest Missouri State College. Currently, Dr. Gies holds the position of Associate Director of the Center for Educational Improvement at UMC. JON S. PADEN received the BS in Education, the MST and PhD degrees from the University of Missouri—Columbia. He has served as a secondary school science and mathematics teacher. During 1967–68 Dr. Paden was an instructor in the UMC School-College Cooperative Physics Program. He most recently was Chairman of the Science Department in the Kamehameha Schools in Honolulu, Hawaii. His special area of interest lies in computer-assisted, individually prescribed instruction. Presently he holds the position of staff coordinator and research specialist in the Center for Educational Improvement at UMC.

SANDRA LOU STEIN, currently Assistant Professor of Education at the University of South Carolina, received her EdD in Educational Psychology and her MEd in Guidance from Northern Illinois University and her BS in Education from The University of Wisconsin. She previously taught at the junior high school level.

The Journal of Educational Research
(Volume 64, September 1970–August 1971)

DEMBAR EDUCATIONAL RESEARCH SERVICES, INC.
Box 1605, Madison, Wisconsin 53701

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT OF W. BENGAL
LIBRARY
CALCUTTA-19

INDEX

Listed Under Articles and Authors

A

Academic Performances of Matched Groups of Community College and 4-Year College Students, Lincoln H. Hall, July–August 1971, 475.

Actual, Ideal, and Expected Role Concepts of Secondary Counselors, Glenn G. Dahlem, January 1971, 205.

Altman, K. I.; T. E. Linton. Operant Conditioning in the Classroom Setting: A Review of the Research, February 1971, 277.

Analysis of Two Hypotheses Concerning the Relationship Between Creativity and Intelligence, An; John S. Dacey; George F. Madaus, January 1971, 213.

Arena, Thomas. Social Maturity in the Prediction of Academic Achievement, September 1970, 21.

Askov, Eunice; Wayne Otto; Warren Askov. A Decade of Research in Handwriting: Progress and Prospect, November 1970, 99.

Assessment of Efforts to Aid Socially Isolated Elementary School Pupils, Merl E. Bonney, April 1971, 359.

B

Beyer, Harold N.; Eugene R. Oetting. Effect of Encouragement and Reassurance on Students' Performance in Remedial English, October 1970, 59.

Bigelow, Gordon S. Field Dependence—Field Independence in 5- to 10-Year-Old Children, May–June 1971, 397.

Bonney, Merl E. Assessment of Efforts to Aid Socially Isolated Elementary School Pupils, April 1971, 359.

Brody, Erness B. A Note on the Validity of the Minnesota Teacher Attitude Inventory (MTAI), October 1970, 67.

C

Caldwell, William E. The Superintendent's Negotiation Role, October 1970, 73.

Carver, Ronald P. A Test of Hypothesized Relationship Between Learning Time and Amount Learned in School Learning, October 1970, 57.

Chang, Sunnyuh Shin; James Raths. The Schools' Contribution to the Cumulating Deficit, February 1971, 272.

Characteristics of Students Who Could Not Take and

Schools Which Did Not Offer Driver Training, Joe E. Shively; William Asher, December 1970, 185.

Characteristics of Students Who Make Accurate and Inaccurate Self-Predictions of College Achievement, Karl E. Keefer, May–June 1971, 401.

Cheating Among College Graduate Students, Charles H. Zastrow, December 1970, 157.

Cohen, Daniel. Differentiating Motivations Underlying Vocational Choice, January 1971, 229.

College Environment: Its Meaning to Academically Successful and Unsuccessful Undergraduates, The; Dennis Earl Nelson, April 1971, 355.

Comparative Achievement of Two Substandard Groups of University Freshmen, Charles B. Keenen, January 1971, 203.

Comparative Study of the Personality Structures of Persistent and Prospective Teachers, A; Richard R. DeBlassie, March 1971, 331.

Comparison of the Effectiveness of Discovery Versus Didactic Methods and Teacher-Guided Versus Independent Procedures in Principle Learning, A; Carol Wiesner, January 1971, 217.

Computer-Assisted Planning and Scheduling of Individualized Programs of Study in Science and Mathematics at the Secondary Level, Sister Marijane Werner, November 1970, 127.

Conceptual Pre-Structuring for Detailed Verbal Passages, Barton B. Proger; Raymond G. Taylor, Jr.; Lester Mann; John M. Coulson; Robert J. Bayuk, Jr., September 1970, 28.

Control of Reinforcement and Academic Achievement, Ralph L. Shaw; Norman P. Uhl, January 1971, 226.

Counselor Education: A Profession on the Move, Thelma M. Urbick; Douglas R. Gross, April 1971, 339.

Credibility of Information for Educational Innovation, Frank B. Murray, September 1970, 17.

Curriculum Development and Evaluation, Hulda Grobman, July–August 1971, 435.

Curriculum Theory, James B. Macdonald, January 1971, 195.

D

Dacey, John S.; George F. Madaus. An Analysis of Two Hypotheses Concerning the Relationship Between Creativity and Intelligence, January 1971, 213.

Dahlem, Glenn G. Actual, Ideal, and Expected Role Concepts of Secondary Counselors, January 1971, 205.

Daines, James R. Test Difficulty as a Factor in Achievement, November 1970, 139.

DeBlassie, Richard R. A Comparative Study of the Personality Structures of Persistent and Prospective Teachers, March 1971, 331.

Decade of Research in Handwriting: Progress and Prospect, A; Eunice Askov; Wayne Otto; Warren Askov, November 1970, 99.

Differentiating Motivations Underlying Vocational Choice, Daniel Cohen, January 1971, 229.

Dobbs, Ralph C. Self-Perceived Solutions to Adult Problems, March 1971, 315.

Durkin, Dolores. What Does Research Say About the Time to Begin Reading Instruction? October 1970, 51.

E

Eberlein, E. Larry. The EPPS Need Structure of In-Service Elementary School Teachers, November 1970, 112.

Effect of Ability Grouping on Concept Learning, The; Joseph C. Piland; Elmer A. Lemke, January 1971, 209.

Effect of Encouragement and Reassurance on Students' Performance in Remedial English, Harold N. Beyer; Eugene R. Oetting, October 1970, 59.

Effect of Grade Contracts on Student Performance, The; William A. Poppen; Charles L. Thompson, May–June 1971, 420.

Effect of Group Counseling on Creativity, S. Jeffrey Garfield; Helen A. Cohen; Robert M. Roth; Harris L. Berenbaum, January 1971, 235.

Effect of Selected Demographic Factors on Prediction of GPA by Male Freshman Students, The; Richard F. Sebok, May–June 1971, 411.

Effect of the System Approach on Changing Preservice Teacher Attitudes Toward Selected Instructional Design Factors, Thomas G. Kennedy; Robert A. Humphrey, April 1971, 365.

Effects of a Mediational Device on Paired-Associate Learning; Learning, and Retention of Sixth Grade Students, The; William E. Loadman, January 1971, 223.

Effects of Certain Audio and Visual Presentation Sequences on Pupil Information Acquisition, Armand J. Galfo, December 1970, 172.

Effects of Same Sex Class Organization on Junior High School Students' Academic Achievement, Self-Discipline, Self-Concept, Sex Role Identification, and Attitude Toward School, Joseph R. Ellis; Joan L. Peterson, July-August 1971, 455.

Effect of Selected Media Feedback Upon the Interactive Behavior of Student Teachers, The: B. Charles Leonard; Frederick John Gies; Jon S. Paden, July-August 1971, 478.

Effects of Training on the Instructional Behaviors of Student Tutors, Fred C. Niedermeyer, November 1970, 119.

Effects of Type of Examination Anticipated on Test Preparation and Performance, The; A. Ralph Hakstian, March 1971, 319.

Ellis, Joseph R.; Joan L. Peterson. Effects of Same Sex Class Organization on Junior High School Students' Academic Achievement, Self-Discipline, Self-Concept, Sex Role Identification, and Attitude Toward School, July–August 1971, 455

Empirical Comparison of Selected Enrollment Projection Formulas, An; William J. Webster, April 1971, 375.

EPPS Need Structure of In-Service Elementary School Teachers, The; E. Larry Eberlein, November 1970, 112.

F

Factors Related to the Effectiveness of Preschool Programs with Disadvantaged Children, Julia R. Vane; William M. Davis, March 1971, 297.

Factors Relating to the Innovativeness of School Systems, Leo R. Hilfiker, September 1970, 23.

Factors Which Are Associated with the Principal's Behavior in Encouraging Teacher Participation in School Decision-Making, Richard A. Gorton, March 1971, 325.

Faculty Allocation of Effort and Student Course Evaluations, Claude W. Grant, May–June 1971, 405.

Field Dependence—Field Independence in 5- to 10-Year-Old Children, Gordon S. Bigelow, May–June 1971, 397.

Fischer, Constance T. Levels of Cheating Under Conditions of Informative Appeal to Honesty, Public Affirmation of Value, and Threats of Punishment, September 1970, 12.

Five-Year Follow-Up of High Ability Achieving and Non-Achieving College Freshmen, A; Leland Smith, January 1971, 220.

Flexibility of Cognitive Style and Its Relationship to Academic Achievement in Fourth and Sixth Grades, Pearline P. Yeatts; Gerald A. Strag, April 1971, 345.

Forecasting the Academic Achievement of Engineering Freshmen, Louis Ross, March 1971, 307.

G

Galfo, Armand J. Effects of Certain Audio and Visual Presentation Sequences on Pupil Information Acquisition, December 1970, 172.

Gamsky, Neal R.; Faye Williams Lloyd. A Longitudinal Study of Visual Perceptual Training and Reading Achievement, July–August 1971, 451.

Garfield, S. Jeffrey; Helen A. Cohen; Robert M. Roth; Harris L. Berenbaum. Effects of Group Counseling on Creativity, January 1971, 235.

Glick, Oren. Interaction Effects of Sex of Mother's Child and Child's Reading Performance on the Mother's Evaluation of the School, November 1970, 124.

Golub, Lester S. A Model for Teaching Composition, November 1970, 115.

Gorton, Richard A. Factors Which are Associated with the Principal's Behavior in Encouraging Teacher Participation in School Decision-Making, March 1971, 325.

Goddard, Alton R. Recency of Secondary School Mathematical Preparation as an Influence on First-Year College Mathematics Achievement, July–August 1971, 443.

Grant, Claude W. Faculty Allocation of Effort and Student Course Evaluations, May–June 1971, 405.

Grobman, Hulda. Curriculum Development and Evaluation, July–August 1971, 435.

H

Haddad, Wadi' D. Nature and Language of Scientific Terms in the Arabic Press of the Middle East, July–August 1971, 465.

Hall, Lincoln H. Academic Performances of Matched Groups of Community College and 4-Year College Students, July–August 1971, 475.

Hakstian, A. Ralph. The Effects of Type of Examination Anticipated on Test Preparation and Performance, March 1971, 319.

Hilfiker, Leo R. Factors Relating to the Innovativeness of School Systems, September 1970, 23.

Howe, Michael J. A. Using Students' Notes to Examine the Role of the Individual Learner in Acquiring Meaningful Subject Matter, October 1970, 61.

I

Implications for Educational Research of Emerging Planning and Management Systems in Higher Education, The; G. Benjamin Lawrence; Clifton C. Lovell; Edward H. Lyell, March 1971, 291.

Instrument for Evaluating Survey Research, An; Richard L. Kohr; Marilyn N. Suydam, October 1970, 78.

Intelligent Word Associations in High School Students, Philip Lambert; Lee H. Hansen; Edgar F. Borgatta, February 1971, 269.

Interaction Effects of Sex of Mother's Child and Child's Reading Performance on the Mother's Evaluation of the School, Oren Glick, November 1970, 124.

Interrelationships Among Self-Esteem, Personal Values, and Interpersonal Values, The; Sandra Lou Stein, July–August 1971, 448.

K

Keefer, Karl E. Characteristics of Students Who Make Accurate and Inaccurate Self-Predictions of College Achievement, May–June 1971, 401.

Keenen, Charles B. Comparative Achievement of Two Substandard Groups of University Freshmen, January 1971, 203.

Kennedy, Thomas G.; Robert A. Humphrey. Effect of the System Approach on Changing Preservice Teacher Attitudes Toward Selected Instructional Design Factors, April 1971, 365.

Khan, Abdul Wahab. Timetable in a Technical University, March 1971, 300.

Kohr, Richard L.; Marilyn N. Suydam. An Instrument for Evaluating Survey Research, October 1970, 78.

L

Lambert, Philip; Lee H. Hansen; Edgar F. Borgatta. Intelligent Word Associations in High School Students, February 1971, 269.

Lawrence, G. Benjamin; Clifton C. Lovell; Edward H. Lyell. The Implications for Educational Research of Emerging Planning and Management Systems in Higher Education, March 1971, 291.

Leonard, B. Charles; Frederick John Gies; Jon S. Paden. The Effect of Selected Media Feedback Upon the Interactive Behavior of Student Teachers, July–August 1971, 478.

Levels of Cheating Under Conditions of Informative Appeal to Honesty, Public Affirmation of Value, and Threats of Punishment, Constance T. Fischer, September 1970, 12.

Loadman, William E. The Effects of a Mediational Device on Paired-Associate Learning; Learning and Retention of Sixth Grade Students, January 1971, 223.

Longitudinal Study Examining Prediction of Doctoral Success: Grade Point Average as Criterion, or Graduation vs. Non-Graduation as Criterion, A; John D. Williams; Steven D. Harlow; Del Gab, December 1970, 161.

Longitudinal Study of Visual Perceptual Training and Reading Achievement, A; Neal R. Gamsky; Faye Williams Lloyd, July–August 1971, 451.

M

Macdonald, James B. Curriculum Theory, January 1971, 195.

McColly, William. What Does Educational Research Say About the Judging of Writing Ability? December 1970, 147.

Mehrens, William A.; Bruce G. Rogers. Relations Between Grade Point Averages and Collegiate Course Grade Distributions, December 1970, 169.

Mental Ability, School Achievement, and Language Arts Achievement in the Prediction of Delinquency, Betty Ruth Raygor, October 1970, 68.

Model for Teaching Composition, A; Lester S. Golub, November 1970, 115.

Murray, Frank B. Credibility of Information for Educational Innovation, September 1970, 17.

N

Nature and Language of Scientific Terms in the Arabic Press of the Middle East, Wadi' D. Haddad, July–August 1971, 465.

Nelson, Deane D. A Study of Personality Adjustment Among Adolescent Children with Working and Nonworking Mothers, March 1971, 328.

Nelson, Dennis Earl. The College Environment: Its Meaning to Academically Successful and Unsuccessful Undergraduates, April 1971, 355.

Niedermeyer, Fred C. Effects of Training on the Instructional Behaviors of Student Tutors, November 1970, 119.

Noffsinger, Thomas; Virginia Dobbs. Teaching Arithmetic to Educable Mentally Retarded Children (Review), December 1970, 177.

Note on the Validity of the Minnesota Teacher Attitude Inventory (MTAI), A; Erness B. Brody, October 1970, 67.

Null, Eldon J. Relationships Between Personal Variables of Teachers and Their Perception of the Behavior of School Personnel, April 1971, 351.

O

Operant Conditioning in the Classroom Setting: A Review of the Research, K. I. Altman; T. E. Linton, February 1971, 277.

Otto, Wayne; Thomas C. Barrett; Richard J. Smith; Kenneth L. Dulin; Dale D. Johnson; Robert Barganz; Susan Hawk; Juanita Peters. Summary and Review of Investigations Relating to Reading, July 1, 1969 to June 30, 1970, February 1971, 242.

P

Perception of Teacher-Educators in India About the Ideal Pupil, T. N. Raina; M. K. Raina, March 1971, 303.

Piland, Joseph C.; Elmer A. Lemke. The Effect of Ability Grouping on Concept Learning, January 1971, 209.

Poppen, William A.; Charles L. Thompson. The Effect of Grade Contracts on Student Performance, May–June 1971, 420.

Proger, Barton B.; Raymond G. Taylor, Jr.; Lester Mann; John M. Coulson; Robert J. Bayuk, Jr. Conceptual Pre-Structuring for Detailed Verbal Passages, September 1970, 28.

R

Raina, T. N.; M. K. Raina. Perception of Teacher-Educators in India About the Ideal Pupil, March 1971, 303.

Raygor, Betty Ruth. Mental Ability, School Achievement, and Language Arts Achievement in the Prediction of Delinquency, October 1970, 68.

Recency of Secondary School Mathematical Preparation as an Influence on First-Year College Mathematics Achievement, Alton R. Goddard, July–August 1971, 443.

Relations Between Grade Point Averages and Collegiate Course Grade Distributions, William A. Mehrens; Bruce G. Rogers, December 1970, 169.

Relationship Between Ego Identity Status and Satisfaction with College, The; Alan S. Waterman; Caroline K. Waterman, December 1970, 165.

Relationships Between Personal Variables of Teachers and Their Perception of the Behavior of School Personnel, Eldon J. Null, April 1971, 351.

Richmond, Bert O.; William F. White. Sociometric Predictors of the Self Concept Among Fifth and Sixth Grade Children, May–June 1971, 425.

Robertson, James R.; John D. Haas. Teacher Personality and the New Social Studies, November 1970, 133.

Rogers, R. S.; E. N. Wright. A Study of Children's Drawings of Their Classrooms, April 1971, 370.

Ross, Louis. Forecasting the Academic Achievement of Engineering Freshmen, March 1971, 307.

S

Schoenfeld, Clay. Toward a National Strategy for Environmental Education, September 1970, 3.

Schools' Contribution to the Cumulating Deficit, The; Sunnyuh Shin Chang; James Raths, February 1971, 272.

Sebok, Richard F. The Effect of Selected Demographic Factors on Prediction of GPA by Male Freshman Students, May–June 1971, 411.

Sekyra, Francis III; M. Ray Loree. Taped Instruction Effects of Problem-Solving Skills of Seventh-Grade Children, October 1970, 86.

Self-Oriented Sentence Completion Responses and Reported Personal Biographical Data, Richard G. Weigel; Fred McKinney, January 1971, 201.

Self-Perceived Solutions to Adult Problems, Ralph C. Dobbs, March 1971, 315.

Sensory-Feedback Modes as Determinants of Learning and Memory, Harvey M. Sussman; Karl U. Smith, October 1970, 64.

Sex Differences in the School Gains of Migrant Children, Sushila Singhal; Priscilla H. Crago, May–June 1971, 417.

Shaw, Ralph L.; Norman P. Uhl. Control of Reinforcement and Academic Achievement, January 1971, 226.

Shively, Joe E.; William Asher. Characteristics of Students Who Could Not Take and Schools Which Did Not Offer Driver Training, December 1970, 185.

Singhal, Sushila; Priscilla H. Crago. Sex Differences in the School Gains of Migrant Children, May–June 1971, 417.

Smith, Leland. A 5-Year Follow-Up Study of High Ability Achieving and Non-Achieving College Freshmen, January 1971, 220.

Social Maturity in the Prediction of Academic Achievement, Thomas Arena, September 1970, 21.

Sociometric Predictors of the Self Concept Among Fifth and Sixth Grade Children, Bert O. Richmond; William F. White, May–June 1971, 425.

State of the Art of Teacher Education, Lindley J. Stiles, May–June 1971, 387.

Stein, Sandra Lou. Interrelationships Among Self-Esteem, Personal Values, and Interpersonal Values, The; July–August 1971, 448.

Stiles, Lindley J. State of the Art of Teacher Education, May–June 1971, 387.

Student Reaction to the Grade Contract, Hugh Taylor, March 1971, 311.

Study Habit Modification and Its Effect on Academic Performance: A Behavioral Approach, Donald J. Tosi; Richard D. Briggs; Rosemary M. Morley, April 1971, 347.

Study of Children's Drawings of Their Classrooms, A; R. S. Rogers; E. N. Wright, April 1971, 370.

Study of Personality Adjustment Among Adolescent Children with Working and Nonworking Mothers, A; Deane D. Nelson, March 1971, 328.

Summary and Review of Investigations Relating to Reading, July 1, 1969 to June 30, 1970, Wayne Otto; Thomas C. Barrett; Richard J. Smith; Kenneth L. Dulin; Dale D. Johnson; Robert Barganz; Susan Hawk; Juanita Peters, February 1971, 242.

Superintendent's Negotiation Role, The; William E. Caldwell, October 1970, 73.

Sussman, Harvey M.; Karl U. Smith. Sensory-Feedback Modes as Determinants of Learning and Memory, October 1970, 64.

T

Taped Instruction Effects on Problem-Solving Skills of Seventh-Grade Children, Francis Sekyra III; M. Ray Loree, October 1970, 86.

Taylor, Hugh. Student Reaction to the Grade Contract, March 1971, 311.

Teacher Personality and the New Social Studies, James R. Robertson; John D. Haas, November 1970, 133.

Teaching Arithmetic to Educable Mentally Retarded Children (Review), Thomas Noffsinger; Virginia Dobbs, December 1970, 177.

Test Difficulty as a Factor in Achievement, James R. Daines, November 1970, 139.

Test of Hypothesized Relationship Between Learning Time and Amount Learned in School Learning, A; Ronald P. Carver, October 1970, 57.

Testing for Number Readiness: Application of the Piagetian Theory of the Child's Development of the Concept of Number, Robert Williams, May–June 1971, 394.

Timetable in a Technical University, Abdul Wahab Khan, March 1971, 300.

Tosi, Donald J.; Richard D. Briggs; Rosemary M. Morley. Study Habit Modification and Its Effect on Academic Performance: A Behavioral Approach, April 1971, 347.

Toward a National Strategy for Environmental Education, Clay Schoenfeld, September 1970, 3.

U

Urbick, Thelma M.; Douglas R. Gross. Counselor Education: A Profession on the Move, April 1971, 339.

Use of the Canonical Correlation Analysis in an Investigation of Pupil's Rate of Learning in School, Margaret C. Wang, September 1970, 35.

Using Students' Notes to Examine the Role of the Individual Learner in Acquiring Meaningful Subject Matter, Michael J. A. Howe, October 1970, 61.

V

Vane, Julia R.; William M. Davis. Factors Related to the Effectiveness of Preschool Programs with Disadvantaged Children, March 1971, 297.

W

Wang, Margaret C. Use of the Canonical Correlation Analysis in an Investigation of Pupil's Rate of Learning in School, September 1970, 35.

Waterman, Alan S.; Caroline K. Waterman. The Relationship Between Ego Identity Status and Satisfaction with College, December 1970, 165.

Webster, William J. An Empirical Comparison of Selected Enrollment Projection Formulas, April 1971, 375.

Weigel, Richard G.; Fred McKinney. Self-Oriented Sentence Completion Responses and Reported Personal Biographical Data, January 1971, 201.

Werner, Sister Marijane. Computer-Assisted Planning and Scheduling of Individualized Programs of Study in Science and Mathematics at the Secondary Level, November 1970, 127.

What Does Educational Research Say About the Judging of Writing Ability? William McColly, December 1970, 147.

What Does Research Say About the Time to Begin Reading Instruction? Dolores Durkin, October 1970, 51.

Wiesner, Carol. A Comparison of the Effectiveness of Discovery Versus Didactic Methods and Teacher-Guided Versus Independent Procedures in Principle Learning, January 1971, 217.

Williams, John D.; Steven D. Harlow; Del Gab. A Longitudinal Study Examining Prediction of Doctoral Success: Grade Point Average as Criterion, or Graduation vs. Non-Graduation as Criterion, December 1970, 161.

Williams, Robert. Testing for Number Readiness: Application of the Piagetian Theory of the Child's Development of the Concept of Number, May–June 1971, 394.

Y

Yeatts, Pearline P.; Gerald A. Strag. Flexibility of Cognitive Style and Its Relationship to Academic Achievement in Fourth and Sixth Grades, April 1971, 345.

Z

Zastrow, Charles H. Cheating Among College Graduate Students, December 1970, 157.

REHABILITATION. INDIA.
CALCUT
70